

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 23:56:24 ; Search time 849.922 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-67
Perfect score: 192
Sequence: 1 GATCCGCGTACTTGCTTTT.....CCGTTACCGAGCCTTCGAGA 192

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ses.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pi.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pin.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	192	100.0	192	6	A68896 Sequence 67
2	192	100.0	192	6	BD062996 DNA and s
3	162.4	84.6	229	1	AF169453 Neisseria
4	162.4	84.6	229	6	A68908 Sequence 79
5	162.4	84.6	229	6	BD063008 DNA and s
6	162.4	84.6	15053	1	AE002550 Neisseria
7	162.4	84.6	172325	6	AX044035 Sequence
8	162.4	84.6	349061	1	NMA222491
9	162.4	84.6	349980	6	AX044034 Sequence
10	160.8	83.8	349	1	AF169440 Neisseria
11	152.8	79.6	400	1	AF169418 Neisseria
12	118.8	61.9	409	1	NMA35415
13	65.4	34.1	1275	6	A96274 Sequence 30
14	65.4	34.1	1572	6	AX024062
15	65.4	34.1	1575	6	A96272 Sequence 30
16	65.4	34.1	1576	6	A96270 Sequence 30
17	60.6	31.6	1575	6	A96276 Sequence 30
18	60.6	31.6	1575	6	AX024113 Sequence
19	36.2	18.9	152572	2	AC090588 Homo sapi
20	36.2	18.9	152881	2	AC025737 Homo sapi
21	36.2	18.9	183169	9	AC009643 Homo sapi
22	35.4	18.4	159539	9	AL512625 Human DNA
23	35.4	18.4	173690	2	AC024055 Homo sapi
24	34.6	18.0	196006	10	AL645973 Mouse DNA
25	34.4	17.9	123004	9	AC087892 Homo sapi
26	34.4	17.9	175733	2	AC016121 Homo sapi
27	34.4	17.9	177162	2	AC068796 Homo sapi
28	34.4	17.9	188361	9	AC087880 Homo sapi
29	34.2	17.8	137705	5	AL929326 Zebrafish
30	34.2	17.8	180594	10	AC112270 Mus muscu
31	34	17.7	202267	9	DJ293M10
32	34	17.7	220895	10	AC125081 Mus muscu
33	33.6	17.5	855239	9	AC026339 Homo sapi
34	33.6	17.5	206030	9	AC021439 Homo sapi
35	33.4	17.4	232297	2	AC097198 Rattus no
36	32.8	17.1	1603	3	FLPAP12
37	32.8	17.1	74586	6	AX356520 Sequence
38	32.8	17.1	74586	6	AX522013 Sequence
39	32.8	17.1	78160	9	AL136102 Human DNA
40	32.8	17.1	143145	9	AL591377 Human DNA
41	32.8	17.1	159817	2	BX322798 Danio rer
42	32.8	17.1	161218	5	AL772148 Zebrafish
43	32.8	17.1	173854	2	AL137023 Homo sapi
44	32.8	17.1	176361	2	BX322785 Danio rer
45	32.8	17.1	192846	2	AC020585 Homo sapi

ALIGNMENTS

RESULT 1
A68896
LOCUS A68896 192 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 67 from Patent WO9802547;
ACCESSION A68896
VERSION A68896.1 GI:4759815
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 192)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Marker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL

```

APPLICATIONS
JOURNAL Patent: WO 9802547-A 67 22-JAN-1998;
COMMENT INST NAT SANTE RECH MED (FR)
FEATURES Other publication FR 2751000 19980116.
source Location/Qualifiers
1..192
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 38 a 51 c 39 g 64 t
ORIGIN

Query Match 100.0%; Score 192; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 2e-49;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGGTGGGCACTCTTCCCGA 60
Db |||||
QY 61 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAACAATATTC 120
Db |||||
QY 61 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAACAATATTC 120
Db |||||
QY 121 GCCTTGGCGTCGGAATGTTCAAGCTAGCTGCATCAGCTAATCAGGTTGCCGTTACC 180
Db |||||
QY 181 GAGCCTTCGAGA 192
Db |||||

RESULT 2
BD062996 122 bp DNA linear PAT 27-AUG-2002
LOCUS DNA and specific proteins or peptides of the Neisseria meningitidis
DEFINITION species bacteria, method for obtaining them and their biological
APPLICATIONS applications.
ACCESSION BD062996
VERSION BD062996.1 GI:22608599
KEYWORDS JP 2001504684-A/58.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 192)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 58 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/58
PD 10-APR-2001
PF 11-JUL-1997 JP 19980505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
PH Key Location/Qualifiers.
1..192
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 38 a 51 c 39 g 64 t
ORIGIN

```

```

Query Match 100.0%; Score 192; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 2e-49;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGGTGGGCACTCTTCCCGA 60
Db |||||
QY 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGGTGGGCACTCTTCCCGA 60
Db |||||
QY 61 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAACAATATTC 120
Db |||||
QY 61 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAACAATATTC 120
Db |||||
QY 121 GCCTTGGCGTCGGAATGTTCAAGCTAGCTGCATCAGCTAATCAGGTTGCCGTTACC 180
Db |||||
QY 181 GAGCCTTCGAGA 192
Db |||||

RESULT 3
AF169453/c 229 bp DNA linear BCT 09-AUG-2000
LOCUS Neisseria meningitidis strain Z2491 clone Cm043 unknown sequence.
DEFINITION Neisseria meningitidis
ACCESSION AF169453
VERSION AF169453.1 GI:9754661
KEYWORDS Neisseria meningitidis
SOURCE Neisseria meningitidis
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 229)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Identification of regions of the chromosome of Neisseria
meningitidis and Neisseria gonorrhoeae which are specific to
pathogenic Neisseriae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES Location/Qualifiers
1..229
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:487"
/clone="Cm043"
BASE COUNT 72 a 48 c 65 g 44 t
ORIGIN

Query Match 84.6%; Score 162.4; DB 1; Length 229;
Best Local Similarity 93.8%; Pred. No. 4.3e-40;
Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGGTGGGCACTCTTCCCGA 60
Db |||||
QY 61 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAACAATATTC 120
Db |||||
QY 144 CATCATCTAAATTTGCTTTTATGTTTACGCCACTCATTCGGGATAAAC-ATATTC 86
QY 121 GCCTTGGCGTCGGAATGTTCAAGCTAGCTGCATCAGCTAATCAGGTTGCCGTTACC 180
Db |||||
QY 85 GCCTTGGCGTCGGAATGTTCAAGCTGCCTGCATCAGCTGCTTACGTTGCCGTTACC 26
QY 181 GAGCCTTCGAGA 192
Db |||||
QY 25 GAGCCTTCGAGA 14

```

RESULT 4
 A68908/c
 LOCUS
 DEFINITION
 A68908
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified
 unclassified
 unclassified
 Nasset, X., Tinsley, C., Achtman, M., Ruelle, J., Vinals, C. and Merker, P.
 'DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL APPLICATIONS'
 Patent: WO 9802547-A 79 22-JAN-1998;
 INST NAT SANTE RECH MED (FR)
 Other publication FR 2751000 19980116.
 FEATURES
 source
 1..229
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 72 a 48 c 65 g 44 t
 ORIGIN
 Query Match 84.6%; Score 162.4; DB 6; Length 229;
 Best Local Similarity 93.8%; Pred. NO. 4.3e-40;
 Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 1 GATCCCGCTACTGGTTTTCATATTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 60
 Db 204 GATCCCGCTACTGGTTTTCATATTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 145
 QY 61 CATCATCTAAATTTGCTTTTATGTTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 120
 Db 144 CATCATCTAAATTTGCTTTTATGTTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 86
 QY 121 GCCTTCCGTCGGATGTTCAAGCTAGCTGCATCACCCTATCAGGTTGCCCTTACC 180
 Db 85 GCCTTCCGTCGGATGTTCAAGCTAGCTGCATCACCCTATCAGGTTGCCCTTACC 26
 QY 181 GAGCCTTCGAGA 192
 Db 25 GAGCCTTCGAGA 14
 RESULT 5
 BD063008/c
 LOCUS
 DEFINITION
 BD063008
 DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.
 BD063008
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified
 unclassified
 unclassified
 Nasset, X., Tinsley, C., Achtman, M., Ruelle, J.L., Vinals, C. and Merker, P.
 'DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL APPLICATIONS'
 Patent: JP 2001504684-A 70 10-APR-2001;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM
 PN JP 2001504684-A/70

PD 10-APR-2001
 PF 11-JUL-1997 JP 19980505685
 PR 12-JUL-1996 FR 96/08768
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI CARLA VINALS,
 PI PETRA MERKER
 PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 1..229
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 72 a 48 c 65 g 44 t
 ORIGIN
 Query Match 84.6%; Score 162.4; DB 6; Length 229;
 Best Local Similarity 93.8%; Pred. NO. 4.3e-40;
 Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 1 GATCCCGCTACTGGTTTTCATATTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 60
 Db 204 GATCCCGCTACTGGTTTTCATATTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 145
 QY 61 CATCATCTAAATTTGCTTTTATGTTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 120
 Db 144 CATCATCTAAATTTGCTTTTATGTTTTCATAGTCTTGTGGTGGGCAATCTTCCCGA 86
 QY 121 GCCTTCCGTCGGATGTTCAAGCTAGCTGCATCACCCTATCAGGTTGCCCTTACC 180
 Db 85 GCCTTCCGTCGGATGTTCAAGCTAGCTGCATCACCCTATCAGGTTGCCCTTACC 26
 QY 181 GAGCCTTCGAGA 192
 Db 25 GAGCCTTCGAGA 14
 RESULT 6
 AE002550/c
 LOCUS
 DEFINITION
 AE002550
 Neisseria meningitidis serogroup B strain MC58 section 192 of 206 of the complete genome.
 AE002550 AE002098
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria meningitidis MC58
 Neisseria meningitidis MC58
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 Neisseriaceae; Neisseria.
 1 (bases 1 to 15053)
 Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Citton, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qian, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V., Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.
 Complete genome sequence of Neisseria meningitidis serogroup B strain MC58
 Science 287 (5459), 1809-1815 (2000)
 JOURNAL
 MEDLINE
 20175755
 PUBMED
 10710307
 REFERENCE
 2 (bases 1 to 15053)
 Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Citton, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Rappuoli, R. and Venter, J.C.
 TITLE
 Complete genome sequence of Neisseria meningitidis serogroup B strain MC58

Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
 Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
 Rappuoli,R. and Venter,J.C.
 Direct Submission
 Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 COMMENT
 On Apr 4, 2000 this sequence version replaced gi:7227263.
 FEATURES
 Location/Qualifiers
 source
 1..15053
 /organism="Neisseria meningitidis MC58"
 /mol_type="genomic DNA"
 /strain="MC58"
 /db_xref="taxon:122586"
 /note="serogroup: B"
 complement(74..982)
 /gene="NMB2000"
 complement(74..982)
 /gene="NMB2000"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42327.1"
 /db_xref="GI:7227264"
 /translation="MQTAINRAVDVTRFIDFDPVRLGHLVLENVQHIWQKXPA
 AIRRALGELLAAVLLSGNLNEGTLIVQVQGRGLKMLVAEASDRTVATARMDET
 AIADDESGLDGLGGGVFLTLQPKDEQWQGVVLEGGAQVQVLMYMKRSQQLDT
 HIVLSASDEAAGLLVQRLPEVLDEAEWHTVSLARTLTAEELAGLDAAQVLYRLPH
 ETPRVFPEPEFSCVTSRCKVSDMLMLGEEVGVVQGSIEVDCCDFCHSKYVF
 DSDYNALFGEDEVGVAKGLPHTVQ"
 1227..1982
 /gene="NMB2001"
 1227..1982
 /gene="NMB2001"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42328.1"
 /db_xref="GI:7227265"
 /translation="WDSFFKPAVAVLWLPFAVRLPALADELTNLLSREQILROFAD
 EEPVLPINRAPARRAGNDELIGSAMGLNEQVLPVNPVRRPARRAGNDELIGSAMGLN
 EQVLPVNPVRRPARRAGNDELIGSAMGLNGLIAYRPGTSTVSTGDCSGFWOHLFKRAM
 GINLPTSAEQRMGTPVARELQDQWVFFRTLGSSRISRVGLYIGNNRFIAPRTG
 KNIETISUSHKYWSGKYAFARRVKNDPSRFLN"
 2001..2219
 /gene="NMB2002"
 2001..2219
 /gene="NMB2002"
 /note="hypothetical protein; identified by Glimmer2;
 putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAF42329.1"
 /db_xref="GI:7227266"
 /translation="MSMPEMPKWDQDVQIVCTEKVKVMSENMAELYQTAQDAFEDA
 LLMGGERQLRDYLLALIEGLENPYRKV"
 2364..2711
 /gene="NMB2003"
 2364..2711
 /gene="NMB2003"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42330.1"
 /db_xref="GI:7227267"
 /translation="MNIIRALLIILGCLATGETAVFLAGIKLPQSVGMGVLPALLQA
 GWKTSMLQQLTDALMSNLTLFLVPPCVAVISYLDLTDADDWFSILVSASATLCLVIV

TKGVRHWIRGIIR"
 2708..3400
 /gene="NMB2004"
 2708..3400
 /gene="NMB2004"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42331.1"
 /db_xref="GI:7227288"
 /translation="MNEILRQPSVLLFLTLAVYALIVTRTGNIFCNPLVLSIVL
 IAYLKILGIDYAVYHNAAQIFDWLKPAAVVLAVPLVQNRKIFNQWLPVIVSQAQS
 VTGIVTGMFYAKWGAEREWLSLASKVTNPPIAIEITRSIGGIPIAITAATVIAAGLV
 GOIAGYKMLKNTVMFSSVMSGLTASHAMGAASLERSRRVAAYAGLGLTENGVLTA
 LIAPLLIPVLGP"
 3467..4687
 /gene="NMB2005"
 3467..4687
 /gene="NMB2005"
 /note="similar to GB:M65216 SP:P38434 GB:S85363 PID:150243
 PID:246444 percent identity: 98.77; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="glutamate N-acetyltransferase/amino-acid
 acetyltransferase"
 /protein_id="AAF42332.1"
 /db_xref="GI:7227269"
 /translation="WAVNLTAKTARQLPDDIGIALYTAQAGVKKPGHTDLTLLIAVAA
 STVGAVFTINRCFAPVHIAKSHLDEDDGVRALVINTGNAGTGAQRIDALAVCAA
 AARQICKPNQVLPFSTGVLPLEPADKIIAALPRMQPAFWNEARAIMIDIVPRAA
 SREGKVDKHTVRTGIAGSGMHPNMTMIGFIATPAKVSQVQVGAAMGAKCKHGLA
 NIITVDGDTSTNDSFVIATGKNSOIEIDNADPRYQAKLLELCLALELQAIIVRDS
 EGATKEITVRVENATRDARQAAVAVARSPLVKTAFASDNLGRLLAAGYAGVAD
 LHTDVEVYLDLILVAHSGRAASTEAQOQAVMSKAEITVRIKLHRGQAATVYTC
 LSHGVYSINADYRS"
 4747..6156
 /gene="NMB2006"
 4747..6156
 /gene="NMB2006"
 /note="similar to PID:1652876 percent identity: 52.04;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="chloride channel protein-related protein"
 /protein_id="AAF42333.1"
 /db_xref="GI:7227270"
 /translation="MTSTFPRRLARKIROTRLRSKSIAPFLLAGSALVALTALFFA
 HADPALELNKLVOQYPMFAVPLPLGLIPLWLRKFPAPFTAGSGIPQVIAISLPL
 YGNAQTRLIRLQCTLLKPLTLFLGMLFGASIGREGPSVQVGAAMGAKCKHGLA
 FQMGENDLMAAGACGLAAAFANAPLAGVIFAIELRGIMLRWERQILLGVLSGFI
 QVAIQNNPYFSGNGVLEHIFLWVSLGVCGAAGGLFGHLLIRGAAAFAPRKIRG
 FVIRNPLLLAALMGLLLALTGTGYQKTYGTGHEAQAHLGIIYEAFFGLAAQWLIAT
 VFSYAGVPGGFTFSLTITGAVLGEHIAAIDISQANIILICMAAFLAGATQSPIT
 SAVVMVPGGQSLIFWMLIACIFASQVSRQSPFPYFVHASGMRFQRVLQETAQTG
 NAPARQQTANSKTPENSEN"
 6346..7737
 /gene="NMB2007"
 /note="this region contains a match to at least one other
 gene that is not full length, and is not the result of a
 sequencing artifact; similar to GP:1742299 percent
 identity: 79.36; identified by sequence similarity;
 putative"
 7881..9524
 /gene="NMB2008"
 7881..9524
 /gene="NMB2008"
 /note="similar to PID:862629 percent identity: 53.03;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11


```
/product="ABC transporter, ATP-binding protein-related
protein".
/protein_id="AAF42334.1"
/db_xref="GI:1722721"
/translations="MDEIQPKKVELQTKLENIKVLKSGSTIIIVGANGTKRLAV
YIEQKFAHRIASRAHLKLNPNVNIPEKSAKTYLSYQWMDGIDVSNRKNYRWDN
NSYTHLNDQWMLQYFAQONNIANNKRNKNEKTKLIDLOEAWETLLPH
RKLHITADIOVASVNEELYGSANMSDGERALFYILGOVLSVDDGSLVPEPELHI
KXSIISLNDKIEELPDGCSFLIITHDIFRATRAVAKXVIIRNYPTPAMDISEVPE
NPBEETITWILGSRKILPVEGNNSLDIATRYCYPDWTIIPKAGCKDVIOSVSLK
KXSEPLNLKSGVLDLSDRDERIEQNLNLGIYILPVESEIENLFSITDVAKLSK
LNGYDEELNLKNGFKSELIYIDNELKDDLDEFVQVRKIDNLKNDLSLKI
TSDMKSLINEISTEOKIETWISIEKNEIQRCIEQDLDKLLTIYDNKGLLAKSA
CVLGMKRNKHEPESWIMRTLKGNKDFIDAIRQKLPILD"
9568..9750
/gene="NMW2009"
/notes="This region contains a gene with one or more
premature stops or frame shifts, and is not the result of a
sequencing artifact; similar to SP:Q09530 PID:733602
percent identity: 61.70; identified by sequence
similarity; putative"
9962..11539

gene
Query Match 84.6%; Score 162.4; DB 1; Length 15053;
Best Local Similarity 93.8%; Pred. No. 5.7e-40;
Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 GATCCGGTACTTGGTTTTCATATTGTCATAGTCTTGCGTGGGCAATCTCCCGA 60
Db 11636 GATCCGGTACTTGGTTTTCATATTGTCATAGTCTTGCGTGGGCAATCTCCCGA 11577

QY 61 CATCATCTAAATTTGCTTTTATGTTTATGCCACTCATTCGGGATAAACAATATCC 120
Db 11576 CATCATCTAAATTTGCTTTTATGTTTATGCCACTCATTCGGGATAAACA-ATATTC 11518

QY 121 GCCTTCGGTCGGATGTTTCAGCTAGCTGCATCACCGTATCATAGTTGCCGTAC 180
Db 11517 GCCTTCGGTCGGATGTTTCAGCTAGCTGCATCACCGTATCATAGTTGCCGTAC 11458

QY 181 GAGCCTTCGAGA 192
Db 11457 GAGCCTTCGAGA 11446

RESULT 7
AX044035/c
LOCUS
DEFINITION
Sequence 114 from Patent WO0066791.
ACCESSION
AX044035
VERSION
AX044035.1 GI:11342919
KEYWORDS
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1
Pizzia,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Maignan,V., Galeotti,C., Mora,M., Ratt,G., Scarcelli,M.,
Scariato,V., Rappelli,R., Frazer,C.M. and Grandi,G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 114 09-NOV-2000;
CHIRON CORPORATION (US); THE INSITUTE FOR GENOMIC RESEARCH (US)
Location/Qualifiers
1..172325
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/notes="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
```

```
BASE COUNT 2272325 172325 bases"
43072 a 47583 c 41465 g 40205 t
ORIGIN
Query Match 84.6%; Score 162.4; DB 6; Length 172325;
Best Local Similarity 93.8%; Pred. No. 6.8e-40;
Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 GATCCGGTACTTGGTTTTCATATTGTCATAGTCTTGCGTGGGCAATCTCCCGA 60
Db 24748 GATCCGGTACTTGGTTTTCATATTGTCATAGTCTTGCGTGGGCAATCTCCCGA 24689

QY 61 CATCATCTAAATTTGCTTTTATGTTTATGCCACTCATTCGGGATAAACAATATCC 120
Db 24688 CATCATCTAAATTTGCTTTTATGTTTATGCCACTCATTCGGGATAAACA-ATATTC 24630

QY 121 GCCTTCGGTCGGGAAATGTTCAAGTAGCTGCATCACCGTAAATCAGTTGCCGTAC 180
Db 24629 GCCTTCGGTCGGGAAATGTTCAAGTAGCTGCATCACCGTAAATCAGTTGCCGTAC 24570

QY 181 GAGCCTTCGAGA 192
Db 24569 GAGCCTTCGAGA 24558

RESULT 8
NMA222491
LOCUS
DEFINITION
Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 2/7.
ACCESSION
AL162753
VERSION
AL162753.2 GI:7379120
KEYWORDS
Neisseria meningitidis Z2491
Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 349061)
Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Rellwell,T., Hamlin,N.,
Holroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skilton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
Nature 404 (6777), 502-506 (2000)
JOURNAL
MEDLINE
20222556
PUBMED
10761319
REFERENCE
2 (bases 1 to 349061)
Parkhill,J.
Direct Submission
Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
Notes:
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N\_meningitidis/).
FEATURES
Location/Qualifiers
1..349061
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/notes="serogroup: A"
complement(24..206)
repeat_unit
310..39"
/label=ATR
209..212
RBS
220..681
gene
/gene="NMA0368"
```

```

CDS
220. .681
/ gene="NMA0368"
/ note="NMA0368, probable integral membrane protein, len: 133 aa; contains four probable transmembrane domains"
/ codon_start=1
/ trans_table=11
/ product="putative integral membrane protein"
/ protein_id="CAB83669.1"
/ db_xref="GI:7379121"
/ db_xref="SPTREMBL:Q9JWH7"
/ translation="WCEQNRKSPFIIVMLLVSVALKIASLSNVAFYLGNGHSMGLTV
LILGIFASLDIRYCAVANYAAVFLAAIIVLLALRKVPVHAAFLVAFSVKAVYV
DAAGNTSDIVRYCAGFYLYWYAAFAVASIGITFAGKNKKAASADGTKNDY"
671. .1492
/ gene="hemK"
/ note="hemK"
671. .1492
/ gene="NMA0369, hemK, HemK protein, len: 273 aa; similar to e.g. HEMK ECOLI P37186 HEMK protein (277 aa), fasta scores; E(): 0, 42.3% identity in 279 aa overlap. Contains P500092 N-6 Adenine-specific DNA methylases signature"
/ codon_start=1
/ trans_table=11
/ product="HemK protein"
/ protein_id="CAB83670.1"
/ db_xref="GI:7379122"
/ db_xref="SPTREMBL:Q9JWH6"
/ translation="WTFDKWLGLSKPKNEARMQLQVSEYTRVOLLTRGGEEMPDEV
RORADLQRRNGEPVAVILGARFEYGRRTVNPVLIIPRTEHLVAVLARLPEN
GRWDLGTGSGAVATVALERPDATVRASDISPPALETARKNADLGRVFAVGSWF
DTDMSEGRWDJISNPFIENGDKHLSQGLRFPFALDIFDGLSCIRILQAQAP
DRLAEGGFLLHGHFDQGAARVGLAENGFSGVETLPDLADLRVTLGKMKHLK"
831. .840
/ gene="hemK"
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
1162. .1171
/ gene="hemK"
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
1187. .1207
/ gene="hemK"
/ note="P500092 N-6 Adenine-specific DNA methylases signature"
complement(1279. .1288)
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
1565. .1568
/ gene="hemK"
1578. .2966
/ gene="NMA0370"
1578. .2966
/ gene="NMA0370"
/ note="NMA0370, probable integral membrane protein, len: 462 aa; similar to hypothetical proteins e.g. Y325_HAEIN P44640 hypothetical protein.Hi0325 (450 aa), fasta scores; E(): 0, 49.8% identity in 464 aa overlap"
/ codon_start=1
/ trans_table=11
/ product="putative integral membrane protein"
/ protein_id="CAB83671.1"
/ db_xref="GI:7379123"
/ db_xref="SPTREMBL:Q9JWH5"
/ translation="NNAVVAVIVMLVLSLRHVHVLSLTIFAFVGGAVAGMPLQNTA
DAAGVSOAGIIPVFNKGELGAKIALSYMGAFAIMATITGSLPQOLAGAVAKLNR
GEMPSVRGSEGVKXLLLSITLVGMQSONIPIHIAFIPMIVPPLLIVNRLKIDR
RLIACVITGLVITFLPIYFGAIFLNEILLGNHSAAPQLDVKNINVMAMAPAL
GVLAGLLAFHYRKFRPYQSNADTAGNADANRPQPSAYRSIAAAVAIVACRAIQ
MYEDSLVLGMLGFAVFMGLVINRDKANDVFGEGIKVMAMVFIMIAAQFAVMA
TGHIOPLVSEMSKMGMAALMVGLLVITMGISGSFSTPLPIAIIYVPLCVGL
GSPATVAIVGTAGALGAGSPASDSTLGTPTMGLNADQDHDRDSVPTFIHYNIP
LLIAGWIAAMVL"
2967. .3191
/ gene="slyX"

```

```

2967. .3191
/ gene="slyX"
/ note="NMA0371, slyX, SLX protein homolog, len: 74 aa; similar to SLX_HAEIN P44759 SLX protein homolog (73 aa), fasta scores; E(): 0.0017, 33.8% identity in 74 aa overlap, and SLX_ECOLI P10857 SLX protein (72 aa), fasta scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/ codon_start=1
/ trans_table=11
/ product="SLX protein homolog"
/ protein_id="CAB83672.1"
/ db_xref="GI:7379124"
/ db_xref="SPTREMBL:Q9JWH4"
/ translation="MDVAQEFHRIETEIQSALQEDVIAGLVAMVAELRQTLDDQQA
QLRLLYQKMDRNPDAQEYFSLRDEIPPHY"
complement(3271. .3423)
/ gene="NMA0372"
complement(3271. .3423)
/ gene="NMA0372"
/ note="NMA0372, unknown, questionable CDS, len: 50 aa"
/ codon_start=1
/ trans_table=11
/ product="very hypothetical protein NMA0372"
/ protein_id="CAB83673.1"
/ db_xref="GI:7379125"
/ db_xref="SPTREMBL:Q9JWH3"
/ translation="MFGRLSPILGSDGFLVAVVEYVGVFCLOGKTAIFRSRGGKRPRI
KGLISD"
3562. .3571
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
complement(3586. .4356)
/ gene="thiF"
complement(3586. .4356)
/ gene="thiF"
/ note="NMA0373, thiF, probable ThiF protein, len: 256 aa; similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa), fasta scores; E(): 0_43.1% identity in 246 aa overlap, and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB protein. (249 aa), fasta scores; E(): 0, 43.9% identity in 244 aa overlap (note that N.m. does not have orthologs of any other molybdopterin biosynthesis proteins). Contains pfam match to entry PF00899 ThiF family, ThiF family"
/ codon_start=1
/ trans_table=11
/ product="ThiF protein"
/ protein_id="CAB83674.1"
/ db_xref="GI:7379126"
/ db_xref="SPTREMBL:Q9JWH2"
/ translation="MTTTEHNDNDAFLRYSRHILDEIGEQQKLSAAHILVVGCG
GLGAALPYLAASGIGTLTIADSDTVLHNLQOVAFDEGVDGKLTALADRLRHIN
HTVDVTINEKLDGCLTGLVOAADIVLCCDNVATROAVNACVQAKTPLYSGAAR
FEQLAVYRDLDFSPCYACLDPGSGASDGICSLFGVFSPLVIGTQAAALAKILL
DAGEPHGRLAVTRALEGGVQVFDLPNPECPVCGAER"
3787. .3796
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
complement(3847. .4257)
/ gene="thiF"
/ note="Fram match to entry PF00899 ThiF family, ThiF family, score 186.60, E-value 4.1e-52"
complement(4418. .4427)
/ note="Core DNA uptake sequence: gccgtctgaa"
/ label=DUS
4471. .7224
/ gene="ppc"
4471. .7224
/ gene="ppc"
/ EC number="4.1.1.31"
/ note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len: 917 aa; similar to many e.g. CAPP RHOP4 O32483 phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa), fasta scores; E(): 0, 43.3% identity in 928 aa overlap.

```

	Query Match	84.6%	Score 12.4	DB 6	Length 349980
	Best Local Similarity	93.8%	Pred. No. 7.2e-40		
	Matches 180	Conservative 0	Mismatches 11	Indels 1	Gaps 1
QY	1	GATCCGCGACTTGGTTTTCATATTTGCATAGCTTGTGCGTCGGCGCATCTTCCCGA	60		
Db	324748	GATCCGCGACTTGGTTTTCATATTTGCATAGCTTGTGCGTCGGCGCATCTTCCCGA			
QY	61	CATCATCTAAATTTGCTTTTATTGTTTTACGCCACCTATTGGCGGATAACAATATTC	120		

```

BASE COUNT      76 a      97 c      85 g      91 t
ORIGIN
Query Match      83.8%; Score 160.8; DB 1; Length 349;
Best Local Similarity 93.2%; Pred. No. 1.4e-39;
Matches 179; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

Qy      1  GATCGCGGTACTTGTTTTTCATATTTGCGATAGCTTTCGGTTCGGGCACTCTCCCGCA 60
      |||
Db      1  GATCGCGGTACTTGTTTTTCATATTTGCGATAGCTTTCGGTTCGGGCTCTCTCCCGCA 60
      |||

Qy     61  CATCATCTAAATTTGCTTTATTGTTTTCACGCCACTCATTCGGGATAAACAATATTC 120
      |||
Db     61  CATCATCTAAATTTGCTTTATTGTTTTCACGCCACTCATTCGGGATAAAC-ATATTC 119
      |||

Qy     121  GCCTTCGCGTCGCGAAATGTTCAAGCTAGCTGCATCACCGTATACAGGTTGCCCGTTAC 180
      |||
Db     120  GCCTTCGCGTCGCGAAATGTTCAAGCTCGCTGCATCACCGGTAATCAGTGTGCCCGTTAC 179
      |||

Qy     181  GAGCCTTCGCGA 192
      |||
Db     180  GAGCCTTCGCGA 191
      |||

```

RESULT 11	
AF169418	
LOCUS	400 bp DNA linear BCT 09-AUG-2000
DEFINITION	Neisseria gonorrhoeae strain FA1090 clone BG007 unknown sequence.
ACCESSION	AF169418
VERSION	AF169418.1 GI:9754626
KEYWORDS	.
SOURCE	Neisseria gonorrhoeae
ORGANISM	Neisseria gonorrhoeae

```

Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
AUTHORS Perrin.A., Nassif.X. and Tinsley.C.R.
TITLE Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 400)
Perrin.A., Nassif.X. and Tinsley.C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156 rue de Vaugirard, Paris 75015, France
FEATURES
Source
1..400
/organism="Neisseria gonorrhoeae"
/mol_type="genomic DNA"
/strains="FA1090"
/db_xref="taxon:485"
/clone="Bg007"
BASE COUNT 87 a 104 c 98 g 111 t
ORIGIN
Query Match 79.6%; Score 152.8; DB 1; Length 400;
Best Local Similarity 90.6%; Pred.No. 4.7e-37;
Matches 174; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
QY 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGCGTGGGCATCTTCCCGGA 60
Db 1 GATCGCGTACTGGTTTTCATATTTGCATAGTCTTGTGCGTGGGCATCTTCCCGGA 60
QY 61 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAACAATATTC 120
Db 61 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAACAATATTC 119
QY 121 GCCTTGCCTCGGATGTTCAAGTACCTGCATCAGCGTTCGCCGTTACC 180
Db 120 GCCTTGCCTCGGATGTTCAAGTACCTGCATCAGCGTTCGCCGTTACC 179
QY 181 GAGCCTTCGAGA 192
Db 180 GAGCCTTCGAGA 191
RESULT 12
NMAJ5415/c
LOCUS 409 bp DNA linear BCT 15-DEC-1998
DEFINITION Neisseria meningitidis partial DNA sequence, clone hrtB.
ACCESSION AJ005415
VERSION AJ005415.1 GI:3212117
KEYWORDS Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
AUTHORS Claus, H., Froesch, M. and Vogel, U.
TITLE Identification of a hotspot for transformation of Neisseria meningitidis by shuttle mutagenesis using signature-tagged transposons
JOURNAL Mol. Gen. Genet. 259 (4), 363-371 (1998)
MEDLINE 99005251
PUBMED 9790590
REFERENCE
AUTHORS 2 (bases 1 to 409)
Vogel, U.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1998) Vogel U., University of Wuerzburg, Institute for Hygiene and Microbiology, Josef-Schneider-Str. 2, 97080 Wuerzburg, GERMANY
COMMENT Partial sequence of meningococcal locus giving rise to high transformation frequency.
FEATURES
Source
1..409
Location/Qualifiers

```

```

/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strains="B1340"
/isolate="German"
/db_xref="taxon:487"
/clone="hrtB"
BASE COUNT 111 a 98 c 105 g 91 t 4 others
ORIGIN
Query Match 61.9%; Score 118.8; DB 1; Length 409;
Best Local Similarity 88.0%; Pred.No. 2.4e-26;
Matches 162; Conservative 0; Mismatches 19; Indels 3; Gaps 3;
QY 10 ACTTGTTTTTCATATTTTCATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCTA 69
Db 391 AATTGGTTTTTCATATTTTCATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCTA 333
QY 70 AATTGCTCTTATTGTTTTTACGCCACTCATTCGGGATAACAATATTCGCCCTTGC 128
Db 332 AATTGCTCTTATTGTTTTTAAAGCCACTCATTCGGGATAAAC-ATATTCGCCCTTGC 274
QY 129 GTCCGAATGTTCAAGCTAGCTGCATCACCGTATCATAGTTGCCCGTTACCGAGCTTC 188
Db 273 ATCCGGAATGCTCAAGCTGCCTGCATCACCGGTAATCAGTTGCCCGTTACCGAGCTTC 214
QY 189 GAGA 192
Db 213 GCGA 210
RESULT 13
A96274/c
LOCUS 1275 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 307 from Patent WO9924578.
ACCESSION A96274
VERSION A96274.1 GI:6780029
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
AUTHORS Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.
TITLE Neisserial antigens
JOURNAL Patent: WO 9924578-A 307 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES
Source
1..1275
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 300 a 304 c 319 g 352 t
ORIGIN
Query Match 34.1%; Score 65.4; DB 6; Length 1275;
Best Local Similarity 87.4%; Pred.No. 1.7e-09;
Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 98 TCATTCCGGATAACAATATTCGCCCTTCGTCGCGAATGTTCAAGCTAGCTGCATCA 157
Db 1275 TCATTCCGGATAAAC-ATATTCGCCCTTCGTCGCGAATGTTCAAGCTAGCTGCATCA 1217
QY 158 CCGTAATCAGTTGCCCGTTACCGAGCTTCGAGA 192
Db 1216 CCGTAATCAGTTGCCCGTTACCGAGCTTCGAGA 1182
RESULT 14
AX024062/c
LOCUS 1572 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 5 from Patent FR2785293.
ACCESSION AX024062
VERSION AX024062.1 GI:10184374

```

KEYWORDS	Neisseria meningitidis
SOURCE	Neisseria meningitidis
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
REFERENCE	Neisseriaceae; Neisseria.
AUTHORS	Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
JOURNAL	Bouchardon, A. and Renaud, M.G.
FEATURES	Patent: FR 2785293-A 5 05-MAY-2000;
source	PASTEUR MERIEUX SERUMS VACC (FR)
CDS	Location/Qualifiers 1..1572 /organism="Neisseria meningitidis" /mol_type="genomic DNA" /db_xref="taxon:487" 1..1572 /note="unnamed protein product" /codon_start=1 /transl_table=11 /protein_id="CAC08907.1" /db_xref="GI:10184375" /translation="MKRSIFVFLYSLLTASEIAYRFVFGIETLPAKMAETPALT MIAALYLFARYKASRLLIIVFAFSAIANNHYAVYQSWMTGNYMLMLEITVEGSA GASMLDKLWLPALWGVLVLEWLCASFKAHKRTHFSADILFLMLMI EVRSFDTKQE H GI SPKPTYSRI KANYESFGYFVGVRVLPYQLDLSRI PAPKOPAPSKI GQGSVONILVI MGSEGAHLKLPGYGRETSPPFLTRLSQADRPDI VKOSV SAGMETAVSLPSFENAIPH ANGLEOISGGDTNWFLAKEQGVEYTVYSAQRNEMAILNLIGKWIDHLDILOPTOLGY GNGDNDEKLLPLFPKINLQGRHPFIVLHQRGSHAPYSALLQPQDKVFGELIVDKDY NTHTKDQMIVTFEGLQKPDPGNLFFATSDHGQVVRQDIYNQGTVPDPSLYLVPLVL YSKNVAQQAAQAQAFAPCEIAFHQLSTELIHTLGVDMPVSQREGSVTGNIITGDAG SLNIRDGAEYVYFPQ"
BASE COUNT	380 a 383 c 395 g 414 t
ORIGIN	
Query Match	34.1%; Score 65.4; DB 6; Length 1572;
Best Local Similarity	87.4%; Pred. No. 1.8e-09;
Matches	83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
Qy	98 TCATTGGCGATAAACAAATATTCGCCCTTCGCCGTAATGTTCAAGCTTAGCCTGCATCA 157
Db	1572 TCATTGGCGATAAAC-AAATTCGCCCTTCGCCGTAATGTTCAAGCTGCCTGCATCAC 1514
Qy	158 CCCTAATCAGGTGCGCGTTACCGAGCCTTCGAGA 192
Db	1513 CCCTAATCAGGTGCGCGTTACCGAGCCTTCGCGA 1479
RESULT 15	
A96272/c	
LOCUS	A96272 1575 bp DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 305 from Patent WO9924578.
ACCESSION	A96272
VERSION	A96272.1 GI:6780028
KEYWORDS	unidentified
SOURCE	unidentified
ORGANISM	unclassified
REFERENCE	1.
AUTHORS	Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.
TITLE	Neisserial antigens
JOURNAL	Patent: WO 9924578-A 305 20-MAY-1999; PIZZA MARIAGRAZIA (IT); SCARLATO VINCENTO (IT); RAPPUOLI RINO (IT); CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES	Location/Qualifiers 1..1575 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644" 376 a _378 c 405 g 416 t
BASE COUNT	376 a _378 c 405 g 416 t
ORIGIN	
Query Match	34.1%; Score 65.4; DB 6; Length 1575;
Best Local Similarity	87.4%; Pred. No. 1.8e-09;

KEYWORDS	Neisseria meningitidis
SOURCE	Neisseria meningitidis
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
REFERENCE	Neisseriaceae; Neisseria.
AUTHORS	Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
JOURNAL	Bouchardon, A. and Renaud, M.G.
FEATURES	Patent: FR 2785293-A 5 05-MAY-2000;
source	PASTEUR MERIEUX SERUMS VACC (FR)
CDS	Location/Qualifiers 1..1572 /organism="Neisseria meningitidis" /mol_type="genomic DNA" /db_xref="taxon:487" 1..1572 /note="unnamed protein product" /codon_start=1 /transl_table=11 /protein_id="CAC08907.1" /db_xref="GI:10184375" /translation="MKRSIFVFLYSLSLTASEIAYRFVFGIETLPAKMAETPALT MIAALYLFARYKASRLLIIVFAFSAIANNHYAVYQSWMTGNYMLMLEITEVGSA GASMLDKLWLPALWGVLVEVLEWLFCSAKFKRTHFSADILFLMLMI EVRSFDTKQE H GI SPKTYTRIGIKANYESFGYFVGRVLPYQLDLSRI PAP KQPAPSKI GQGSGVONILV MGSEGAHLKLPGYGRETSPPFLTSLQADRPDI VKOSV SAGMTAVSLPSFENAIPH ANGLEOISGGDTNWFLAKEQGVETYFYSAQENEMAILNLIGKWIDHLLIOPTOLGY GNGDNDEKLLPLFPKINLQGRHPFIVLHQRGSHAPYSALLPOQDKVFGELIVDKDY NTHKIDQMIVTFEGLOKQOPGNLFFATSDHGQVRQDIYNQGTVPDPSLYLVPLVL YSKNVAQQAAQAQAFAPCEIAFHQLSTELIHTLGVDMPVSQREGSVTGNIITGDAG SLNIRDGKAERYVFPQ"
BASE COUNT	380 a 383 c 395 g 414 t
ORIGIN	
Query Match	34.1%; Score 65.4; DB 6; Length 1572;
Best Local Similarity	87.4%; Pred. No. 1.8e-09;
Matches	83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
Qy	98 TCATTGGCGATAAACAATATTCGCCCTTCGCCGTAATGTTCAAGCTGCATCA 157
Db	1572 TCATTGGCGATAAAC-ATATTCGCCCTTCGCCGTAATGTTCAAGCTGCATCAC 1514
Qy	158 CCCTAATCAGGTGCGCGTTACCGAGCCTTCGAGA 192
Db	1513 CCCTAATCAGGTGCGCGTTACCGAGCCTTCGCGA 1479
RESULT 15	
A96272/c	
LOCUS	A96272 1575 bp DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 305 from Patent WO9924578.
ACCESSION	A96272
VERSION	A96272.1 GI:6780028
KEYWORDS	unidentified
SOURCE	unidentified
ORGANISM	unclassified
REFERENCE	1.
AUTHORS	Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.
TITLE	Neisserial antigens
JOURNAL	Patent: WO 9924578-A 305 20-MAY-1999; PIZZA MARIAGRAZIA (IT); SCARLATO VINCENTO (IT); RAPPUOLI RINO (IT); CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES	Location/Qualifiers 1..1575 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:3264" 376 a _378 c 405 g 416 t
BASE COUNT	376 a _378 c 405 g 416 t
ORIGIN	
Query Match	34.1%; Score 65.4; DB 6; Length 1575;
Best Local Similarity	87.4%; Pred. No. 1.8e-09;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 123.917 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-67

Perfect score: 192

Sequence: 1 GATCCGCGTACTTGCTTTT.....CGGTACCGAGCCTTCGAGA 192

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	192	19	AAV03576
2	162.4	84.6	229	19	AAV03587
3	162.4	84.6	12438	21	AA081550
4	162.4	84.6	172325	21	AAF21613
5	162.4	84.6	349980	21	AAF21612
6	162.4	84.6	837096	21	AA081489
7	65.4	34.1	366	21	AA081358
8	65.4	34.1	624	20	AAZ12104
9	65.4	34.1	1275	20	AAZ12106
10	65.4	34.1	1572	21	AAZ15298
11	65.4	34.1	1575	20	AAZ12105
12	60.6	31.6	1575	20	AAZ12107
13	60.6	31.6	1575	21	AAZ15323
14	59.2	30.8	1467	25	ABZ41628
15	35.4	18.4	12100	22	AAZ40550
16	35.4	18.4	12100	22	AAZ06438
17	35.4	18.4	12100	22	AAZ69502
18	35.4	18.4	12100	22	AAZ72028
19	32.8	17.1	74586	24	AAZ72705
20	32.4	16.9	21737	23	ABZ16905
21	32.4	16.9	23253	23	ABZ06944
22	32.4	16.9	23253	23	ABZ06900
23	32.2	16.8	1599	25	ABZ39988
24	32	16.7	485	25	ABZ52607
25	32	16.7	1201	24	ABZ29400
26	32	16.7	1201	24	ABZ29401
27	32	16.7	3228	25	ABZ41663
28	31.6	16.5	585	25	ABZ54003
29	31.6	16.5	4041	23	ABZ28959
30	31.4	16.4	6781	23	ABZ28958
31	31	16.1	396	24	ABZ77589
32	31	16.1	760	24	ABZ18364
33	31	16.1	760	24	ABZ18365
34	30.8	16.0	678	24	ABZ38658
35	30.8	16.0	678	24	ABZ38663
36	30.8	16.0	3394	25	AAZ51325
37	30.6	15.9	4936	24	AAZ19028
38	30.4	15.8	3185	23	ABZ21362
39	30.4	15.8	97662	22	AAZ83908
40	30	15.6	197	24	ABZ79418
41	30	15.6	7022	23	ABZ22288
42	30	15.6	15161	24	ABZ70457
43	30	15.6	15161	24	AAZ61422
44	29.8	15.5	5046	23	ABZ02755
45	29.8	15.5	17369	23	ABZ02754

ALIGNMENTS

RESULT 1

AAV03576

ID AAV03576 standard; DNA; 192 BP.

AC AAV03576;

XX 22-OCT-1998 (first entry)

DT Neisseria meningitidis DNA sequence B13.

DE N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
meningitis; ss.

OS Neisseria meningitidis.

PN WO9802547-A2.

XX 22-JAN-1998.

PF 11-JUL-1997; 97WO-FR01295.

PR 12-JUL-1996; 96FR-0008768.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (SMIK) SMITHKLINE BEECHAM.

PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI Vinals C;

Neisseria meningit
DNA encoding a pol
Neisseria meningit
Neisseria gonorrh
DNA encoding a pol
N. gonorrhoeae nuc
DNA encoding human
Human reproductive
Human immune/haema
Human immune/haema
Genomic DNA encodi
Drosophila melanog
Drosophila melanog
N. gonorrhoeae nuc
Aspergillus oryzae
Oligonucleotide fo
Oligonucleotide fo
N. gonorrhoeae nuc
Aspergillus oryzae
Drosophila melanog
Drosophila melanog
Human ORF2636 cDNA
Oligonucleotide fo
Oligonucleotide fo
Oligonucleotide fo
Human receptor cDN
Maize Ramosa 1 cod
Drosophila melanog
Genomic sequence o
Bacillus clausii g
Drosophila melanog
Chemically treated
Human gene regulat
Drosophila melanog
Drosophila melanog

DR WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX

PS Example 4; Page 113; 150pp; French.

XX

XX AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, *frpA* or *C*,

CC *opc*, *porA*, *rotamase*, sequence IC1106, *IGA* protease, *pilA*, *pilC*,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC *lactamica*, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX

SQ Sequence 192 BP; 38 A; 51 C; 39 G; 64 T; 0 other;

Query Match 100.0%; Score 192; DB 19; Length 192;

Best Local Similarity 100.0%; Pred. No. 5e-56;

Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCGCGTACTGGTTTTCATATTTTCATAGTCTTTCGTCGGCGCATCTTCCCGGA 60

DB 1 GATCCGCGTACTGGTTTTCATATTTTCATAGTCTTTCGTCGGCGCATCTTCCCGGA 60

QY 61 CATCATCTAAATTTGCTTTTATTTGTTTTCATGCCACCTATTGGCGGATAAACAATATTC 120

DB 61 CATCATCTAAATTTGCTTTTATTTGTTTTCATGCCACCTATTGGCGGATAAACAATATTC 120

QY 121 GCCTTGCCGTCGCGAATGTTCAAGCTAGCTGATCAGCTTCCGTAATCAGTTGCCGTTACC 180

DB 121 GCCTTGCCGTCGCGAATGTTCAAGCTAGCTGATCAGCTTCCGTAATCAGTTGCCGTTACC 180

QY 181 GAGCCTTCGAGA 192

DB 181 GAGCCTTCGAGA 192

RESULT 2

AAV03587/c

ID AAV03587 standard; DNA; 229 BP.

XX AAV03587;

AC

XX

XX

DT 22-OCT-1998 (first entry)

DE *Neisseria meningitidis* DNA sequence C43.

DE

DE

KW *N. gonorrhoeae*; *N. lactamica*; chromosome 22491; region 1; region 2;

KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;

KW meningitis; ss.

XX

OS *Neisseria meningitidis*.

XX

XX WO9802547-A2.

XX

PD 22-JAN-1998.

XX

XX 11-JUL-1997; 97WO-FR01295.

XX

XX 12-JUL-1996; 96FR-0008768.

PR

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (SMIK) SMITHKLINE BEECHAM.

XX

XX Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;

PI Vinals C;

XX

DR WPI; 1998-110594/10.

XX

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX

PS Example 4; Page 119; 150pp; French.

XX

XX AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, *frpA* or *C*,

CC *opc*, *porA*, *rotamase*, sequence IC1106, *IGA* protease, *pilA*, *pilC*,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC *lactamica*, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX

SQ Sequence 229 BP; 72 A; 48 C; 65 G; 44 T; 0 other;

Query Match 84.6%; Score 162.4; DB 19; Length 229;

Best Local Similarity 93.8%; Pred. No. 8.1e-46;

Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 GATCCGCGTACTGGTTTTCATATTTTCATAGTCTTTCGTCGGCGCATCTTCCCGGA 60

DB 204 GATCCGCGTACTGGTTTTCATATTTTCATAGTCTTTCGTCGGCGCATCTTCCCGGA 145

QY 61 CATCATCTAAATTTGCTTTTATTTGTTTTCATGCCACCTATTGGCGGATAAACAATATTC 120

DB 144 CATCATCTAAATTTGCTTTTATTTGTTTTCATGCCACCTATTGGCGGATAAACAATATTC 86

QY 121 GCCTTGCCGTCGCGAATGTTCAAGCTAGCTGATCAGCTTCCGTAATCAGTTGCCGTTACC 180

DB 85 GCCTTGCCGTCGCGAATGTTCAAGCTAGCTGATCAGCTTCCGTAATCAGTTGCCGTTACC 26

QY 181 GAGCCTTCGAGA 192

DB 25 GAGCCTTCGAGA 14

RESULT 3

AAA81550/c

ID AAA81550 standard; DNA; 12438 BP.

XX AAA81550;

AC

XX

XX 04-DEC-2000 (first entry)

XX

XX *N. meningitidis* partial DNA sequence gnm_97 SEQ ID NO:97.

DE

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

KW *Meningococcus B*; MenB; ds.

XX

OS *Neisseria meningitidis*.

XX

XX WO200022430-A2.

XX

XX 20-APR-2000.

PD

XX 08-OCT-1999; 99WO-US23573.

PF


```

XX 09-OCT-1998; 99US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,
PI Rappuoli R, Pizza M,
XX WPI; 2000-318079/27.
DR
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea -
XX
XX Claim 7; Page 1489-1492; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences, AAA81453 to AAA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences, AAA81260 to AAA81303 and AAA82620 to AAA82663 represent
XX Neisseria DNA sequences and their corresponding proteins; AAA81254 to
XX AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
XX sequences, which are all used in the exemplification of the present
XX invention. The nucleic acid sequences, protein sequences, and antibodies
XX against them, can be used in the manufacture of a composition. The
XX composition can be used as a medicament (or in the manufacture of a
XX medicament) for treating, preventing or diagnosing infection due to
XX Neisserial bacteria. For example, some of the identified proteins could
XX be components of vaccines against Meningococcus B; against all serotypes;
XX and/or against all pathogenic Neissariae. Identification of sequences
XX from the bacterium will also facilitate production of biological probes,
XX particularly organism-specific probes. Attempts to make efficacious
XX Meningococcus B vaccines have failed mainly due to antigen tolerance.
XX Multivalent vaccines have also been tried but none have successfully
XX overcome antigenic variability. The provision of further, complete
XX sequences may provide an opportunity to identify secreted or surface
XX exposed proteins that may be presumed targets for the immune system and
XX which are not antigenically variable or at least more conserved than
XX other more variable regions.
XX
XX Sequence 12438 BP; 3661 A; 3104 C; 2565 G; 3108 T; 0 other;
SQ
Query Match 84.6%; Score 162.4; DB 21; Length 12438;
Best Local Similarity 93.8%; Pred No. 4e-45; Mismatches 11; Indels 1; Gaps 1;
Matches 130; Conservative 0;
QY 1 GATCCGCGTACTGGTTTTCATATTTTGCATAGTCTTGTGCGTCGGGCGATCTCCCGGA 60
DB 4568 GATCCGCGTACTGGTTTTCATATTTTGCATAGTCTTGTGCGTCGGGCGATCTCCCGGA 4509
QY 61 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAACATATTCC 120
DB 4508 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAAC-ATATTCC 4450
QY 121 GCCTTGCCGTCGCGAATGTTTCAAGCTAGCTGATCAGCTAATCAGGTTGCCCGTTACC 180
DB 4449 GCCTTGCCGTCGCGAATGTTTCAAGCTGCTGTCATCAGGTAATCAGGTTGCCCGTTACC 4390
QY 181 GAGCCTTCGAGA 192
DB 4389 GAGCCTTCGCGA 4378
RESULT 4
AAF21613/c
ID AAF21613 standard; DNA; 172325 BP.
XX
XX AAF21613;
XX

```

```

DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
OS Neisseria meningitidis.
PN WO200066791-A1.
XX
XX 09-NOV-2000.
XX
XX 08-MAR-2000; 2000WO-US05928.
XX
XX 30-APR-1999; 99US-0132068.
XX 08-OCT-1999; 99WO-US233573.
XX 28-FEB-2000; 2000GB-0004695.
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V,
XX Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R,
XX Frazer CM, Grandi G,
XX WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -
XX Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
XX Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
XX to AAF21613 represent fragments of the NMB genomic sequence, as the
XX sequence was too long to go in a record on its own it was split into 8
XX sequences which overlap each other at the beginning and end of each
XX sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
XX the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
XX the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
XX Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
XX AAF21606 represent PCR primers which are used in the exemplification of
XX the present invention. The NMB genome and fragments from it have
XX antibacterial activity, and can be used in vaccines and gene therapy.
XX Neisseria nucleic acids, proteins and/or antibodies which binds to the
XX proteins can be used in compositions for treating or preventing infection
XX due to Neisserial bacteria or as a diagnostic reagent for detecting the
XX presence of Neisserial bacteria or of antibodies raised to Neisserial
XX bacteria. Computers, computer memory, computer storage medium or computer
XX databases can be used in a search to identify open reading frames (ORFs)
XX or coding sequences within the NMB genome. The DNA sequences provide
XX further opportunities to find antigenic or immunogenic proteins which are
XX more effective in vaccines than the outer membrane proteins currently
XX used.
XX
XX Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;
SQ
Query Match 84.6%; Score 162.4; DB 21; Length 172325;
Best Local Similarity 93.8%; Pred No. 1.1e-44;
Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 1 GATCCGCGTACTGGTTTTCATATTTTGCATAGTCTTGTGCGTCGGGCGATCTCCCGGA 60
DB 24748 GATCCGCGTACTGGTTTTCATATTTTGCATAGTCTTGTGCGTCGGGCGATCTCCCGGA 24689
QY 61 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAACATATTCC 120
DB 24688 CATCATCTAAATTTGCTTTTATTTGTTTACGCCACTCATTCGGGATAAC-ATATTCC 24630
QY 121 GCCTTGCCGTCGCGAATGTTTCAAGCTAGCTGATCAGCTAATCAGGTTGCCCGTTACC 180

```

Db 24629 GCCTTCCGTCGCGAATGTTCAAGCTGCTGCATCACCGGTAATCAGGTGCCGTTACC 24570
 QY 181 GAGCCTTCGAGA 192
 Db 24569 GAGCCTTCGCGA 24558

RESULT 5
 ID AAF21612/c
 XX AAF21612 standard; DNA; 349980 BP.
 AC AAF21612;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KW ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05928.
 XX
 PR 30-APR-1999; 99US-0132068.
 PR 08-OCT-1999; 99WO-US23573.
 PR 28-FEB-2000; 2000GB-0004695.
 XX
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
 PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
 PI Frazer CM, Grandi G;
 XX
 XX WPI; 2000-647603/62.
 DR
 XX Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections -
 PT
 PS Claim 7; Appendix A; 692pp; English.
 XX
 CC The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8
 CC sequences which overlap each other at the beginning and end of each
 CC sequence by 4980 bp (i.e. the last 4980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 4980 bp of AAF21607 are repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the other membrane proteins currently
 CC used.
 XX
 SQ Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;
 Query Match 84.6%; Score 162.4; DB 21; Length 349980;

Best Local Similarity 93.8%; Pred. No. 1.5e-44;
 Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 GATCCGCGTACTGGTTTTTCATATTTTCATAGTCTTTCGTCGGGCGATCTTCCCCGA 60
 Db 324748 GATCCGCGTACTGGTTTTTCATATTTTCATAGTCTTTCGTCGGGCGATCTTCCCCGA 324589

QY 61 CATCATCTAAATTTGCTTTTATTTGTTTTTACGCCACTCATTCGGGATAAACAATATTC 120
 Db 324688 CATCATCTAAATTTGCTTTTATTTGTTTTTACGCCACTCATTCGGGATAAACA-ATATTC 324630

QY 121 GCCTTCCGTCGCGAATGTTCAAGCTGCTGCATCACCGTATCAGGTGCCGTTACC 180
 Db 324629 GCCTTCCGTCGCGAATGTTCAAGCTGCTGCATCACCGTATCAGGTGCCGTTACC 324570

QY 181 GAGCCTTCGAGA 192
 Db 324569 GAGCCTTCGCGA 324558

RESULT 6
 ID AAA81489/c
 XX AAA81489 standard; DNA; 837096 BP.
 AC AAA81489;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO2000022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23573.
 XX
 PR 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 XX (CHIR) CHIRON CORP.
 PA
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX
 XX WPI; 2000-318079/27.
 XX
 PT Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX
 PS Claim 7; Page 629-865; 1760pp; English.
 XX
 CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could

XX The present invention describes methods of obtaining immunogenic
CC proteins from *Neisseria* genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed *Neisseria meningitidis* genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81331 represent PCR primers used in the
CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81332 to
CC AAA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC *Neisseria* bacteria. For example, some of the identified proteins could
CC be components of vaccines against *Meningococcus B*; against all serotypes;
CC and/or against all pathogenic *Neisseriae*. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC *Meningococcus B* vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
XX
SQ Sequence 366 BP; 84 A; 100 C; 97 G; 85 T; 0 other;

Query Match 34.1%; Score 65.4; DB 21; Length 366;
Best Local Similarity 87.4%; Pred. No. 2.2e-12;
Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 98 TCATTCCGGATAACCAATATTCGCTTCGCCGCGAATGTTCAAGCTAGCTGCATCA 157
Db 366 TCATTCCGGATAAAC-ATATTCGCTTCGCCGCGAATGTTCAAGCTAGCTGCATCA 308

QY 158 CCCTAATCAGTTGCCCGTTACCGAGCCTTCGAGA 192
Db 307 CCGTATCAGTTGCCGTTACCGAGCCTTCGCGA 273

RESULT 8
AAZ12104/C
ID AAZ12104 standard; DNA; 624 BP.
XX
XX AAZ12104;
XX
XX
DT 08-OCT-1999 (first entry)
DE
DE *Neisseria meningitidis* partial ORF81 sequence.
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW treatment; *Neisseria* infection; meningitis; septicemia, gonorrhea; ss.
XX
XX *Neisseria meningitidis*.
XX
XX Key Location/Qualifiers
FH misc_feature 256..257
FT /*tag= a
FT /*note= 952 unspecified nucleotides are present
FT between these bases"
XX
XX MO9924578-A2.
XX
XX 20-MAY-1999.
XX
XX 09-OCT-1998; 98MO-IB01565.
XX
XX 01-SEP-1998; 98GB-0019016.
XX 06-NOV-1997; 97GB-0023516.
XX 14-NOV-1997; 97GB-0024190.
XX 18-NOV-1997; 97GB-0024386.
XX

```

PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX (CHIR-) CHIRON SPA.
XX Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX WPI; 1999-327407/27.
DR P-PSDB; AAY38650.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
XX Claim 9; Page 204; 524pp; English.
XX
XX Nucleotide sequences AA211972-212358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
XX Sequence 624 BP; 141 A; 152 C; 159 G; 170 T; 2 other;
SQ
Query Match 34.1%; Score 65.4; DB 20; Length 624;
Best Local Similarity 87.4%; Pred. No. 2.7e-12;
Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 98 TCATTCCGGATAAACAAATATTCGGCTTCGGCTCCGGAATGTTCAAGTAGCCTGCATCA 157
DB 624 TCATTCCGGATAAAC-ATATTCGGCTTCGGCTCCGGAATGTTCAAGCTGCTGCATCAC 566
QY 158 CCGTAATCAGTTGCCCGTTACCGAGCCTTCGAGA 192
DB 565 CCGTAATCAGTTGCCCGTTACCGAGCCTTCGCGA 531
RESULT 9
AA212106/c
ID AA212106 standard; DNA; 1275 BP.
XX
XX AA212106;
XX
XX 08-OCT-1999 (first entry)
XX
XX Neisseria meningitidis strain A complete ORF81 sequence.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
XX
XX Neisseria meningitidis.
XX
XX WO924578-A2.
XX
XX 20-MAY-1999.
XX
XX 09-OCT-1998; 98WO-IB01665.
XX
XX 01-SEP-1998; 98GB-0019016.
XX 06-NOV-1997; 97GB-0023516.
XX 14-NOV-1997; 97GB-0024190.
XX 18-NOV-1997; 97GB-0024386.
XX 27-NOV-1997; 97GB-0025158.
XX 10-DEC-1997; 97GB-0026147.
XX 14-JAN-1998; 98GB-0000759.
XX (CHIR-) CHIRON SPA.
XX
XX Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX

```

```

DR WPI; 1999-327407/27.
DR P-PSDB; AAY38652.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
XX Claim 9; Page 206; 524pp; English.
XX
XX Nucleotide sequences AA211972-212358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
XX Sequence 1275 BP; 300 A; 304 C; 319 G; 352 T; 0 other;
SQ
Query Match 34.1%; Score 65.4; DB 20; Length 1275;
Best Local Similarity 87.4%; Pred. No. 3.6e-12;
Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 98 TCATTCCGGATAAACAAATATTCGGCTTCGGCTCCGGAATGTTCAAGTAGCCTGCATCA 157
DB 1275 TCATTCCGGATAAAC-ATATTCGGCTTCGGCTCCGGAATGTTCAAGCTGCTGCATCAC 1217
QY 158 CCGTAATCAGTTGCCCGTTACCGAGCCTTCGAGA 192
DB 1216 CCGTAATCAGTTGCCCGTTACCGAGCCTTCGCGA 1182
RESULT 10
AAAL5298/c
ID AAAL5298 standard; DNA; 1572 BP.
XX
XX AAAL5298;
XX
XX 04-SEP-2000 (first entry)
XX
XX DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
XX Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
XX Neisseria meningitidis.
XX
XX Key Location/Qualifiers
XX CDS 1..1572
XX /*tag= a
XX
XX WO200026375-A2.
XX
XX 11-MAY-2000.
XX
XX 28-OCT-1999; 99WO-FR02643.
XX
XX 30-OCT-1998; 98FR-0013693.
XX
XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Aujame L, Bouchardon A, Renaud-Mongenie G, Rokbi B, Nassif X;
XX Tinsley C, Perrin A;
XX
XX WPI; 2000-355622/31.
XX P-PSDB; AAY93267.
XX
XX New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX or preventative vaccines and for diagnosis
XX
XX Claim 2; Page 56-59; 187pp; French.
XX
XX The present sequence encodes a protein that is specific for pathogenic
XX

```

CC strains of *Neisseria*. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 XX
 XX
 SQ Sequence 1572 BP; 380 A; 383 C; 395 G; 414 T; 0 other;

Query Match 34.1%; Score 65.4; DB 21; Length 1572;
 Best Local Similarity 87.4%; Pred. No. 3.9e-12;
 Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 98 TCATTGGGATAAACAAATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCA 157
 Db 1572 TCATTGGGATAAAC-ATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCA 1514
 QY 158 CCGTAATCAGGTGCCCGTTACCGAGCCTTCGAGA 192
 Db 1513 CCGTAATCAGGTGCCCGTTACCGAGCCTTCGAGA 1479

RESULT 11
 ID AA212105/C standard; DNA; 1575 BP.
 XX AC AA212105;
 XX DT 08-OCT-1999 (first entry)
 XX DE *Neisseria meningitidis* complete ORF81 sequence.
 XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 XX treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.
 XX OS *Neisseria meningitidis*.
 XX PN WO9924578-A2.
 XX XX
 XX PD 20-MAY-1999.
 XX PF 09-OCT-1998; 98WO-IB01665.
 XX PR 01-SEP-1998; 98GB-0019016.
 XX PR 06-NOV-1997; 97GB-0023516.
 XX PR 14-NOV-1997; 97GB-0024190.
 XX PR 18-NOV-1997; 97GB-0024386.
 XX PR 27-NOV-1997; 97GB-0025158.
 XX PR 10-DEC-1997; 97GB-0026147.
 XX PR 14-JAN-1998; 98GB-0000759.
 XX PA (CHIR-) CHIRON SPA.
 XX PI Grandi G, Massignani V, Pizza M, Rappuoli R, Scarlato V;
 XX DR WPI; 1999-327407/27.
 XX DR P-PSDB; AAY38651.
 XX PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 XX diagnosis, treatment and prevention of infection
 XX PS Claim 9; Page 205; 524pp; English.
 XX CC Nucleotide sequences AA211972-721358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX SQ Sequence 1575 BP; 376 A; 378 C; 405 G; 416 T; 0 other;

Query Match 34.1%; Score 65.4; DB 20; Length 1575;
 Best Local Similarity 87.4%; Pred. No. 3.9e-12;
 Matches 83; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 98 TCATTGGGATAAACAAATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCA 157
 Db 1575 TCATTGGGATAAAC-ATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCA 1517
 QY 158 CCGTAATCAGGTGCCCGTTACCGAGCCTTCGAGA 192
 Db 1516 CCGTAATCAGGTGCCCGTTACCGAGCCTTCGAGA 1482

RESULT 12
 ID AA212107 standard; DNA; 1575 BP.
 XX AC AA212107;
 XX DT 08-OCT-1999 (first entry)
 XX DE *Neisseria gonorrhoeae* complete ORF81 sequence.
 XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 XX treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.
 XX OS *Neisseria gonorrhoeae*.
 XX PN WO9924578-A2.
 XX XX
 XX PD 20-MAY-1999.
 XX PF 09-OCT-1998; 98WO-IB01665.
 XX PR 01-SEP-1998; 98GB-0019016.
 XX PR 06-NOV-1997; 97GB-0023516.
 XX PR 14-NOV-1997; 97GB-0024190.
 XX PR 18-NOV-1997; 97GB-0024386.
 XX PR 27-NOV-1997; 97GB-0025158.
 XX PR 10-DEC-1997; 97GB-0026147.
 XX PR 14-JAN-1998; 98GB-0000759.
 XX PA (CHIR-) CHIRON SPA.
 XX PI Grandi G, Massignani V, Pizza M, Rappuoli R, Scarlato V;
 XX DR WPI; 1999-327407/27.
 XX DR P-PSDB; AAY38653.
 XX PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 XX diagnosis, treatment and prevention of infection
 XX PS Claim 9; Page 208; 524pp; English.
 XX CC Nucleotide sequences AA211972-721358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX SQ Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;

Query Match 31.6%; Score 60.6; DB 20; Length 1575;
 Best Local Similarity 84.2%; Pred. No. 1.7e-10;
 Matches 80; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
 QY 98 TCATTGGGATAAACAAATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCA 157
 Db 1575 TTATTGGGATAAAC-ATATTCGCGCTTCGCGAATGTTCAAGCTAGCTGCATCGC 1517

QY 158 CCGTAATCAGGTTGCCGCTTACCGAGCCTTCGAGA 192
 DB 1516 CCGTAATCAGGTTGCCGCTTACCGAGCCTTCGCGA 1482

RESULT 13

AA15323/C
 ID AAA15323 standard; DNA; 1575 BP.

AC AAA15323;
 AC 04-SEP-2000 (first entry)
 DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.
 DE Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 KW *Neisseria* gonorrhoeae.
 OS
 XX
 XX Key Location/Qualifiers
 XX CDS 1..1575
 XX /*tag= a

PN WO200026375-A2.
 XX
 XX 11-MAY-2000.
 XX
 XX 28-OCT-1999; 99WO-FR02643.
 XX
 XX 30-OCT-1998; 98FR-0013693.
 XX
 XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PI Aujane L, Bouchardon A, Renauld-Mongenie G, Rokhi B, Nassif X;
 PI Tinsley C, Perrin A;
 DR WPI; 2000-365622/31.
 DR P-PSDB; AA93293.
 XX
 XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 PT or preventative vaccines and for diagnosis

PS Claim 4; Page 137-139; 187pp; French.

XX The present sequence encodes a protein that is specific for pathogenic
 CC strains of *Neisseria*. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.

XX Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;

Query Match 31.6%; Score 60.6; DB 21; Length 1575;
 Best Local Similarity 84.2%; Pred. No. 1.7e-10;
 Matches 80; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 98 TCATTGGGATAACAATATTCGCCCTTCCGGAATGTTCAAGCTAGCCTGCATCA 157
 DB 1575 TTATTCCGGATAAAC-ATATTCGCCCTTCCGGAATGTTCAAGCTAGCCTGCATCGC 1517

QY 158 CCGTAATCAGGTTGCCGCTTACCGAGCCTTCGAGA 192
 DB 1516 CCGTAATCAGGTTGCCGCTTACCGAGCCTTCGCGA 1482

RESULT 14

AB241628/C
 ID AB241628 standard; DNA; 1467 BP.

XX

AC AB241628;
 XX
 XX 07-MAR-2003 (first entry)
 DE N. gonorrhoeae nucleotide sequence SEQ ID 7845.
 XX Antibacterial; infection; vaccine; gene therapy; gene; ds.
 KW *Neisseria* gonorrhoeae.
 OS
 XX WO200279243-A2.
 PN
 XX 10-OCT-2002.
 XX
 XX 12-FEB-2002; 2002WO-IB02069.
 XX
 XX 12-FEB-2001; 2001GB-0003424.
 PR
 XX (CHIR-) CHIRON SPA.
 XX
 XX Fontana MR, Pizza M, Masignani V, Monaci E;
 XX
 XX WPI: 2003-058415/05.
 DR P-PSDB; ABP80658.
 DR
 XX New protein from *Neisseria* gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing *N. gonorrhoeae* infection -
 PT
 XX Disclosure; Page 760-761; 815pp; English.
 PS
 XX The present invention relates to proteins from *Neisseria* gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records AB237706-AB242016 represent nucleic acid
 CC molecules of the invention.
 XX
 XX SQ Sequence 1467 BP; 356 A; 358 C; 377 G; 376 T; 0 other;

Query Match 30.8%; Score 59.2; DB 25; Length 1467;
 Best Local Similarity 84.8%; Pred. No. 5.1e-10;
 Matches 78; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 101 TTCCGGATAACAATATTCGCCCTTCCGGAATGTTCAAGCTAGCCTGCATCACCG 160
 DB 1467 TTCCGGATAAAC-ATATTCGCCCTTCCGGAATGTTCAAGCTAGCCTGCATCACCG 1409

QY 161 TAATCAGGTTGCCGCTTACCGAGCCTTCGAGA 192

DB 1408 TAATCAGGTTGCCGCTTACCGAGCCTTCGAGA 1377

RESULT 15

AA40550/C
 ID AA40550 standard; DNA; 12100 BP.

XX
 AC AA40550;
 XX

DT 17-DEC-2001 (first entry)
 DE DNA encoding human prostate cancer antigen, Seq ID No 702.

XX Human; prostate cancer antigen; cytostatic; uropathic; diagnostic;
 KW reproductive system; chromosomal marker; forensic; urinary disorder;
 KW chronic nephritis; blood-related disorder; thrombosis; ds.

OS Homo sapiens.

XX
 XX WO200155316-A2.
 XX

PD 02-AUG-2001.

XX	17-JAN-2001; 2001WO-US01328.	PR	29-SEP-2000; 2000US-0236369.
XX		PR	29-SEP-2000; 2000US-0236370.
XX		PR	02-OCT-2000; 2000US-0236802.
PR	31-JAN-2000; 2000US-0179065.	PR	02-OCT-2000; 2000US-0237037.
PR	04-FEB-2000; 2000US-0180628.	PR	02-OCT-2000; 2000US-0237038.
PR	24-FEB-2000; 2000US-0184664.	PR	02-OCT-2000; 2000US-0237039.
PR	02-MAR-2000; 2000US-01866350.	PR	02-OCT-2000; 2000US-0237040.
PR	16-MAR-2000; 2000US-0189874.	PR	13-OCT-2000; 2000US-0239935.
PR	17-MAR-2000; 2000US-0190076.	PR	13-OCT-2000; 2000US-0239937.
PR	18-APR-2000; 2000US-0198023.	PR	20-OCT-2000; 2000US-0240960.
PR	19-MAY-2000; 2000US-0205515.	PR	20-OCT-2000; 2000US-0241221.
PR	07-JUN-2000; 2000US-0209467.	PR	20-OCT-2000; 2000US-0241785.
PR	28-JUN-2000; 2000US-0214886.	PR	20-OCT-2000; 2000US-0241786.
PR	30-JUN-2000; 2000US-0215135.	PR	20-OCT-2000; 2000US-0241787.
PR	07-JUL-2000; 2000US-0216647.	PR	20-OCT-2000; 2000US-0241808.
PR	07-JUL-2000; 2000US-0216680.	PR	20-OCT-2000; 2000US-0241809.
PR	11-JUL-2000; 2000US-0217487.	PR	20-OCT-2000; 2000US-0241826.
PR	11-JUL-2000; 2000US-0217496.	PR	01-NOV-2000; 2000US-0244617.
PR	14-JUL-2000; 2000US-0218290.	PR	01-NOV-2000; 2000US-0244617.
PR	26-JUL-2000; 2000US-0220963.	PR	08-NOV-2000; 2000US-0246474.
PR	26-JUL-2000; 2000US-0220964.	PR	08-NOV-2000; 2000US-0246475.
PR	14-AUG-2000; 2000US-0224518.	PR	08-NOV-2000; 2000US-0246476.
PR	14-AUG-2000; 2000US-0224519.	PR	08-NOV-2000; 2000US-0246477.
PR	14-AUG-2000; 2000US-0225119.	PR	08-NOV-2000; 2000US-0246478.
PR	14-AUG-2000; 2000US-0225113.	PR	08-NOV-2000; 2000US-0246523.
PR	14-AUG-2000; 2000US-0225166.	PR	08-NOV-2000; 2000US-0246524.
PR	14-AUG-2000; 2000US-0225267.	PR	08-NOV-2000; 2000US-0246525.
PR	14-AUG-2000; 2000US-0225268.	PR	08-NOV-2000; 2000US-0246526.
PR	14-AUG-2000; 2000US-0225270.	PR	08-NOV-2000; 2000US-0246527.
PR	14-AUG-2000; 2000US-0225447.	PR	08-NOV-2000; 2000US-0246528.
PR	14-AUG-2000; 2000US-0225457.	PR	08-NOV-2000; 2000US-0246532.
PR	14-AUG-2000; 2000US-0225758.	PR	08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225759.	PR	08-NOV-2000; 2000US-0246610.
PR	22-AUG-2000; 2000US-0226681.	PR	08-NOV-2000; 2000US-0246611.
PR	22-AUG-2000; 2000US-0226868.	PR	17-NOV-2000; 2000US-0246207.
PR	22-AUG-2000; 2000US-0227182.	PR	17-NOV-2000; 2000US-0246207.
PR	23-AUG-2000; 2000US-0227009.	PR	17-NOV-2000; 2000US-0246208.
PR	30-AUG-2000; 2000US-0228924.	PR	17-NOV-2000; 2000US-0246210.
PR	01-SEP-2000; 2000US-0229287.	PR	17-NOV-2000; 2000US-0246211.
PR	01-SEP-2000; 2000US-0229343.	PR	17-NOV-2000; 2000US-0249212.
PR	01-SEP-2000; 2000US-0229344.	PR	17-NOV-2000; 2000US-0249213.
PR	01-SEP-2000; 2000US-0229345.	PR	17-NOV-2000; 2000US-0249214.
PR	03-SEP-2000; 2000US-0229345.	PR	17-NOV-2000; 2000US-0249215.
PR	03-SEP-2000; 2000US-0229513.	PR	17-NOV-2000; 2000US-0249216.
PR	06-SEP-2000; 2000US-0230437.	PR	17-NOV-2000; 2000US-0249217.
PR	06-SEP-2000; 2000US-0230438.	PR	17-NOV-2000; 2000US-0249218.
PR	08-SEP-2000; 2000US-0231242.	PR	17-NOV-2000; 2000US-0249244.
PR	08-SEP-2000; 2000US-0231243.	PR	17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231244.	PR	17-NOV-2000; 2000US-0249264.
PR	08-SEP-2000; 2000US-0231133.	PR	17-NOV-2000; 2000US-0249265.
PR	08-SEP-2000; 2000US-0231414.	PR	17-NOV-2000; 2000US-0249297.
PR	08-SEP-2000; 2000US-0232080.	PR	17-NOV-2000; 2000US-0249299.
PR	12-SEP-2000; 2000US-0232081.	PR	17-NOV-2000; 2000US-0249300.
PR	14-SEP-2000; 2000US-0232397.	PR	01-DEC-2000; 2000US-0250160.
PR	14-SEP-2000; 2000US-0232398.	PR	01-DEC-2000

PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the reproductive system including prostate cancer
PT and also for testing and detection e.g. diagnosis -
XX
PS
XX Disclosure; SEQ ID No 702; 546pp; English.
XX
CC The invention relates to novel isolated human prostate cancer antigen
CC polynucleotides (I) and polypeptides (II). (I) and (II) are useful for
CC preventing, treating or ameliorating a medical condition when
CC administered. (I), (II) and the antibody to (II) are useful for treating,
CC preventing and/or prognosing disorders related to the reproductive
CC system including prostate cancers; urinary disorders e.g. chronic
CC nephritis; and blood-related disorders e.g. thrombosis. (II) can be used
CC for testing and detection e.g. as a chromosomal marker and in forensics.
CC (I) and the anti-(II) antibody can be used in testing and detection in
CC immunoassays. AAS40061-AAS4075 represent the human prostate cancer
CC antigen coding sequences, and related PCR primers and sequences of the
CC invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12100 BP; 3682 A; 2710 C; 2748 G; 2960 T; 0 other;

Query Match 18.4%; Score 35.4; DB 22; Length 12100;
Best Local Similarity 51.6%; Pred. No. 0.18;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 GATCGCGTACTGGTTTTCATATTTTGCATAGTCTTGTCGGTGGGCATCTTCCCGA 60
Db 4343 GCTCGTAGTACTGGTTGTTATTTTCTCTGGGTCAGTGTACATCCCTT 4284

QY 61 CATCATCTAAATTCGTTTATTTGGTTTTCAGCCACTCATTCGGGATAAACAATATTC 120
Db 4283 CATCATTTCTAATTTCTTTATTGAATCTTCTCTCACTCTTTATTATTAATCTAGTAGT 4224

QY 121 GCCTTGCCTCGCGAATGTTCAAGCTAGCCGATCA 157
Db 4223 GCCCTATCTTGTAAATTTTTCAAAAAACCTTGATTA 4187

Search completed: November 15, 2003, 00:35:20
Job time : 125.917 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1280.85 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-67

Perfect score: 192

Sequence: 1 GATCCGGTACTTGCTTTT.....CGTTACCGAGCCTTCGAGA 192

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152338056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	36.6	19.1	281	10	BF870084 IL3-ET011
C 2	35.4	18.4	1117	9	AL575232 AL575232
C 3	34.2	17.8	429	28	A2499482
C 4	34	17.7	710	28	BH975720

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

PUBMED

10737800

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

ALIGNMENTS

RESULT 1
BF870084/c BF870084 281 bp mRNA linear EST 17-JAN-2001
LOCUS IL3-ET0116-281000-309-B12 ET0116 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BF870084
VERSION BF870084.1 GI:12260214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Ngai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Cosca, F.F., Goldman, G.H., Carvalho, A.F., Matsumura, A., Bala, G.S., Simpson, D.H., Brundstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

5	33.8	17.6	659	28	AQ621897	AQ621897 HS 3107 B
C 6	33.8	17.6	674	28	AZ571385	AZ571385 286PvH04
C 7	33.2	17.3	1044	13	EX436571	EX436571 BX436571
8	33.3	17.2	1060	13	EX436571	EX436571 BX436571
9	32.8	17.1	342	9	AL727786	AL727786 AL727786
10	32.6	17.0	841	13	BQ792412	BQ792412 EST 8867
C 11	32.4	16.9	608	9	AL729504	AL729504 AL729504
12	32.4	16.9	733	9	AL716475	AL716475 AL716475
C 13	32.4	16.9	863	29	CC143842	CC143842 NDL.27P10
C 14	32.2	16.8	935	28	B10635	B10635 F18123-T7 I
C 15	32	16.7	519	9	AA631450	AA631450 np85e02.s
C 16	32	16.7	598	10	AW971216	AW971216 EST383304
C 17	31.8	16.6	937	28	BH137534	BH137534 ENTNA83TF
C 18	31.8	16.6	449	28	BH827894	BH827894 BACP26-J
C 19	31.8	16.6	543	28	BH186641	BH186641 031 G.12-
C 20	31.8	16.6	543	29	CNS07R2P	AL623587 T3 end of
C 21	31.8	16.6	594	12	BP021327	BP021327 BP021327
C 22	31.8	16.6	679	13	BW149556	BW149556 BW149556
C 23	31.8	16.6	716	29	CC157466	CC157466 1519H09.b
C 24	31.8	16.6	732	13	BW017928	BW017928 BW017928
C 25	31.8	16.6	732	13	BW307894	BW307894 BW307894
C 26	31.8	16.6	763	13	BW180122	BW180122 BW180122
C 27	31.8	16.6	792	13	BW084442	BW084442 BW084442
C 28	31.8	16.6	801	13	BW186373	BW186373 BW186373
C 29	31.8	16.6	1034	13	EX433669	EX433669 EX433669
C 30	31.6	16.5	445	9	AA820407	AA820407 LD24006.5
C 31	31.6	16.5	446	9	AA820402	AA820402 LD23994.5
C 32	31.6	16.5	624	28	AZ994418	AZ994418 2M0279E19
C 33	31.6	16.5	632	9	AI542901	AI542901 SD09414.5
C 34	31.6	16.5	657	28	BH984640	BH984640 oe172h06
C 35	31.6	16.5	682	10	BG637449	BG637449 SD15454.5
C 36	31.6	16.5	828	29	CC359511	CC359511 PUH1833TD
C 37	31.6	16.5	878	29	BZ392669	BZ392669 EINBM96TR
C 38	31.4	16.4	507	9	AL882926	AL882926 AL882926
C 39	31.4	16.4	615	9	AL793669	AL793669 AL793669
C 40	31.4	16.4	679	9	AL884840	AL884840 AL884840
C 41	31.4	16.4	743	29	AG045973	AG045973 Pan trogl
C 42	31.4	16.4	943	28	AZ206473	AZ206473 SP_0104.B
C 43	31.4	16.4	1101	29	CNS00218	AL081925 Drosophila
C 44	31.4	16.4	1101	29	CNS00262	AL037268 Drosophila
C 45	31.2	16.2	561	29	EX200869	EX200869 Danio rer

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (GI:4732114|GB|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 141 a 72 c 95 g 121 t
ORIGIN

Query Match 17.8%; Score 34.2; DB 28; Length 429;
Best Local Similarity 62.1%; Pred. No. 1e+02;
Matches 54; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 32 TAGTCTGTGGTGGGCATCTCCCGACATCATCTAAATTTGCTTTTATGGTTTAA 91
|||||
Db 162 TCGTCTTTTGTCCAGCGAGTTTCATGACATATCTATGTTGGCTAGACTGCTTTA 221
|||||

QY 92 CGCCATCATTTGGGATAACAATAT 118
|||||

Db 222 TACCACCACCTAGCATACATACACAGT 248
|||||

RESULT 4
BH975720
LOCUS
DEFINITION
B975720.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 710)
REFERENCE
AUTHORS
Delebaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash
W., Rabinowicz,P.D. and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: sub85 row: a column: 02
Seq primer: -21UPOT forward
Class: shotgun
High quality sequence start: 12
High quality sequence stop: 551.
Location/Qualifiers
1..710
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: POTW13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TOL00DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome

BASE COUNT 104 a 136 c 69 g 350 t
ORIGIN

Query Match 17.6%; Score 33.8; DB 28; Length 659;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 50; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 TTTTTCATATTTGCATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCTAAATTTG 75
|||||

Db 495 TTTTTCCTGTCGCTTAACCTTCTTTGTGGGCTTATCCCTCCCTTCCCTTTAAATTT 554
|||||

BASE COUNT 219 a 102 c 104 g 285 t
ORIGIN

Query Match 17.7%; Score 34; DB 28; Length 710;
Best Local Similarity 57.5%; Pred. No. 1.1e+02;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 11 CTTCGTTTTCATATTTGCATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCTAA 70
|||||

Db 477 CTCAGGTGGCATCTTTTGTACTTTCTTCTGATTTAAATCTTCCAGTACACTTCTGT 536
|||||

QY 71 ATTTGCTTTTATTTGGTTTACGCCACTCATTTGGGATAACATA 116
|||||

Db 537 TTTTCTTTTCTTTTGTGACGTCCTTCTGTTTGTATATAAAAA 582
|||||

RESULT 5
AO621897
LOCUS
DEFINITION
HS_3107.B1.H07.MR.CIT.Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3107 Col=13 Row=P, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AO621897.1 GI:5084289
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 659)
REFERENCE
AUTHORS
Mahaitas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Mahaitas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 3107 row: P column: 13
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 659.
Location/Qualifiers
1..659
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3107 Col=13 Row=P"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelBAC11; BAC clones in E-Coli DH10B"

BASE COUNT 104 a 136 c 69 g 350 t
ORIGIN

Query Match 17.6%; Score 33.8; DB 28; Length 659;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 50; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 TTTTTCATATTTGCATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCTAAATTTG 75
|||||

Db 495 TTTTTCCTGTCGCTTAACCTTCTTTGTGGGCTTATCCCTCCCTTCCCTTTAAATTT 554
|||||


```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1060)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
CGI-Bin/Cluster.cgi?seq=CS0BAH016ZF01NP1&cluster=7394.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAH016ZF01NP1.
Location/Qualifiers
BASE COUNT 175 a 254 c 293 g 221 t 117 others
ORIGIN
1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1082YK09"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 175 a 254 c 293 g 221 t 117 others
ORIGIN
Query Match 17.2%; Score 33; DB 13; Length 1060;
Best Local Similarity 34.2%; Pred. No. 1.9e+02;
Matches 53; Conservative 35; Mismatches 67; Indels 0; Gaps 0;
QY 28 TGCATAGCTTGTGCGTGGCGATCTCCCGCATCATCTAAATTTCTTTTATGTT 87
D 862 TKCAATKVMKXDWGTCGRCTCTCMTCMCYMTTCYTKTTCACAKKTAKGT 921
QY 88 TTATGCCACTCATTCGGGTAACAATATTCGCCGTCGCGGTAATTTCAAGCTA 147
D 922 TTATCMMAGSCGYYTTTTTTTATTTCTTTCGATACSGGGVMMGARTAMAGA 981
QY 148 GCCTCATCACGTAATCAGGTGCGCGTTACCGA 182
D 982 CCKKCSVTAGSVTYGGTTTGKCCCGKGSBA 1016
RESULT 9
AL727786
LOCUS
DEFINITION
AL727786 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BN0AA091ZAI1 5', mRNA sequence.
ACCESSION
AL727786.1 GI:20192390
VERSION
AL727786.1
KEYWORDS
EST.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 342)
AUTHORS
Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M., Hardelin
, J.P., Weissenbach, J. and Petit, C.
TITLE
A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL
Unpublished
CONTACT
Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

```

```

source
1..342
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA091ZAI1"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cDNA"
BASE COUNT 79 a 94 c 75 g 94 t
ORIGIN
Query Match 17.1%; Score 32.8; DB 9; Length 342;
Best Local Similarity 49.4%; Pred. No. 2.4e+02;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 2 ATCCGGGTACTTGGTTTTCATATTTGATAGTCTTTGCGTGGCGATCTTCCCGGAC 61
D 13 ACCCGGCGAGGTACTCTCGATGATTTTCAGCAGGTCTGATTTCCGAACGCTCACAATGA 72
QY 62 ATCATCTAAATTTGTCTTTATTTGTTTACGCCACTTCATTCGGGATAAACAATATTCG 121
D 73 TTCTGATCAGTGTGTTATCATCGTTCCTCCAAACCTTCATAGCATGTTGGAGCTTTCCG 132
QY 122 CCTTGGCGTGGCAATGTTCAAGCTAGCTGCTGATCACCCTTAATCAGTTGCC 173
D 133 CAAAGTAGGCAGGTTTATTTCCAGCAGCTTCCACGAGATCATCAGGCAGTC 184
RESULT 10
BQ792412
LOCUS
DEFINITION
BQ792412 841 bp mRNA linear EST 30-JUL-2002
vinifera cDNA clone CT00SG02 3', mRNA sequence.
ACCESSION
BQ792412
VERSION
BQ792412.1 GI:22007378
KEYWORDS
EST.
SOURCE
Vitis vinifera
ORGANISM
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
1 (bases 1 to 841)
REFERENCE
Abbal, P., Agasse, A., Ageorges, A., Atanassova, R., Barrieu, F.,
Couture, C., Dedaldecamp, F., Delrot, S., Glissant, D., Grimplet, J.,
Hamdi, S., Romieu, C. and Terrier, N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages
JOURNAL
Unpublished
COMMENT
Contract: Hamdi S.
UMR 619 - Equipe Biologie de la Vigne
Universite de Bordeaux I, Institut National de la Recherche
Agronomique
71, Avenue Edouard Bourleaux, BP 81, 33883 Villenave D'Ornon Cedex,
France
Tel: 00-33-(0)5-57-12-25-50
Fax: 00-33-(0)5-57-12-25-48
Email: s.hamdi@bordeaux.inra.fr
Seq primer: T7.
Location/Qualifiers
FEATURES
1..841
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CT00SG02"
/dev_stage="veraison stage"
/clone_lib="Veraison Grape berries Lambda Triplex2
Library"
/note="Organ: Fruit without seeds; Vector: Lambda Triplex2
; Site1: SfiIA; Site2: SfiIB; Oriented library"
BASE COUNT 157 a 225 c 116 g 343 t

```



```

Email: ents@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..863
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="NDL.27P10"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pECBAC1; Site 1: Hind III; The library was
prepared from whole body tissue of newly hatched 1st larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"
219 a 174 c 162 g 308 t
BASE COUNT
ORIGIN
Query Match 16.9%; Score 32.4; DB 29; Length 863;
Best Local Similarity 54.1%; Pred. No. 2.8e+02;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 16 TTTTTCATTTTGGCATAGCTTCGCGTGGGATCTCCCGACATCATCTAAATTG 75
Db 318 TTGTACTAGTTGTGCCACTGATTGTTGTAACATTTTCTGTTCTCTACGATT 377
QY 76 TCTTATTGGTTTTACGCCACTATTGGGATTAACAATATTCGGCTTGGCGTCGGA 135
Db 378 CTCATTTCGATTCGTTTCCACTGATTGCTGTAACACTTTTCTGTTTCTTCACGT 437
QY 136 AT 137
Db 438 AT 439

RESULT 14
B10635/c
LOCUS
DEFINITION
B10635 936 bp DNA linear GSS 14-MAY-1997
F18123-T7 IGF Arabidopsis thaliana genomic clone F18123, genomic
survey sequence.
B10635
B10635.1 GI:2091756
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids
; eurosids. II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 936)
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
Ecker, J.
BAC End Sequences at ATGC
Unpublished
Other_GSSs: F18123-Sp6
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-998-9384
Fax: 215-998-8780
Email: jacker@atgenome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 154
High quality sequence stop: 352.
FEATURES
Location/Qualifiers
1..936
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"

/clone="F18123"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altman"
293 a 233 c 162 g 203 t 45 others
BASE COUNT
ORIGIN
Query Match 16.8%; Score 32.2; DB 28; Length 936;
Best Local Similarity 54.5%; Pred. No. 3.2e+02;
Matches 61; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 22 ATATTTCATAGTCTTCGCGTGGGATCTCCCGACATCATCTAAATTCTCTTA 81
Db 755 ATATTTTTTTTTTTAAAGTTTGGGAGGATTCAGGTGAAAAACCTTAAGTTATTTT 696
QY 82 TTGGTTTTACGCCACTATTGGGATTAACAATATTCGGCTTGGCGTCG 133
Db 695 TTGATTTTAAACCATTCGTGGAGGATCATCAAGGTGACCTTCTGTGCC 644

RESULT 15
AA631450/c
LOCUS
DEFINITION
AA631450 519 bp mRNA linear EST 30-OCT-1997
np85e02.s1 NCI CGAP Thy1 Homo sapiens cDNA clone IMAGE:1133114
similar to gb:Y00282 RIBOPHORIN II PRECURSOR (HUMAN);, mRNA
sequence.
AA631450
AA631450.1 GI:2554061
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnli.gov/bbrp/image/image.html
Insert Length: 603 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 342.
FEATURES
Location/Qualifiers
1..519
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1133114"
/tissue="thyroid"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Thy1"
/note="Vector: pAMP10; mRNA made from invasive thyroid
tumor, cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."
175 a 92 c 111 g 141 t
BASE COUNT
ORIGIN
Query Match 16.7%; Score 32; DB 9; Length 519;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 80; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 29.9056 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-67

Perfect score: 192

Sequence: 1 GATCGCGTACTTGTTT.....CCGTACGACCTTCGAGA 192

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PTCUTUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29.2	15.2	2868	4	US-08-780-562-8
2	29.2	15.2	3097	2	US-08-599-455B-1
3	29.2	15.2	3097	3	US-09-069-781B-1
4	29.2	15.2	3097	4	US-09-137-132-1
5	29.2	15.2	3097	4	US-08-864-564A-1
6	29.2	15.2	3097	4	US-09-094-410-1
7	29.2	15.2	3097	4	US-08-708-123D-1
8	29.2	15.2	3097	4	US-08-583-153A-1
9	29.2	15.2	3097	4	US-08-570-142D-1
10	29.2	15.2	3097	4	US-08-638-524B-1
11	29.2	15.2	3495	3	US-08-827-962-17
12	29.2	15.2	3650	3	US-08-837-635-5
13	29.2	15.2	3650	3	US-08-837-635-7
14	29.2	15.2	3650	3	US-08-827-962-16
15	29.2	15.2	3650	3	US-08-803-345-2
16	29.2	15.2	3854	2	US-08-599-455B-42
17	29.2	15.2	3854	3	US-09-069-781B-42
18	29.2	15.2	3854	4	US-09-137-132-42
19	29.2	15.2	3854	4	US-08-864-564A-42
20	29.2	15.2	3854	4	US-09-094-410-42
21	29.2	15.2	3854	4	US-08-708-123D-42
22	29.2	15.2	3854	4	US-08-638-524B-42
C 23	29.2	15.2	3854	2	US-09-047-026A-3
C 24	29.2	15.2	6854	1	US-08-468-036-4
C 25	29.2	15.2	6854	2	US-08-376-843-4
26	28.8	15.0	7218	1	US-08-232-463-14
C 27	28.8	15.0	11887	4	US-08-961-527-146

Sequence 100, App
Sequence 1, Appli
Sequence 77, Appl
Sequence 565, App
Sequence 519, App
Sequence 493, App
Sequence 609, App
Sequence 695, App
Sequence 2, Appli
Sequence 1, Appli
Sequence 22, Appli
Sequence 17, Appl
Sequence 35, Appl
Sequence 151, App
Sequence 151, App
Sequence 5, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-08-780-562-8
; Sequence 8, Application US/08780562
; Patent No. 6541604
; GENERAL INFORMATION:
; APPLICANT: Matthews, William
; APPLICANT: Bennett, Brian
; TITLE OF INVENTION: WSX RECEPTOR
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,562
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585005
; FILING DATE: 01/08/97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/
; FILING DATE: 01/08/97
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0986R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2868 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-780-562-8

Query Match 15.2%; Score 29.2; DB 4; Length 2868;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;


```

; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
US-09-137-132-1

Query Match      15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred.No.2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY      65  ATCTAAATTGCTTTATTGTTTACGGCACTTCATTCGGGATAACAATATTCGCGCT 124
DB      267  AGCTAAATTAACTCAAGTGGTATCTACGTTCTTGAGTTATCCAAAACAGTCTTCCACTG 326

QY      125  TGCCGTGGCAATGTTCAAGCT 146
DB      327  TTGCTTTGGCAATGAGCAAGGT 348

RESULT 5
US-08-864-564A-1
; Sequence 1, Application US/08864564A
; Patent No. 6395498
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Cupepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,564A
; FILING DATE: 28-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906

```

```
TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
US-08-864-564A-1

Query Match 15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTGCTTTATTGGTTTACGCCACTCATTCGGGATAAACAATATTCGGCT 124
Db 267 AGCTAAATTTAATTCGAAGTGTATCTACGTTCTCGATATCCAAAACAGTCTTCCACTG 326

QY 125 TGCCGTCGCGAATGTTCAAGCT 146
Db 327 TTGCTTTGGGAATGAGCAAGGT 348

RESULT 6
US-09-094-410-1
; Sequence 1, Application US/09094410
; Patent No. 6483552
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,410
; FILING DATE: 09-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/864,564
; FILING DATE: 28-MAY-1997
; APPLICATION NUMBER: 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
US-09-094-410-1

Query Match 15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTGCTTTATTGGTTTACGCCACTCATTCGGGATAAACAATATTCGGCT 124
Db 267 AGCTAAATTTAATTCGAAGTGTATCTACGTTCTCGATATCCAAAACAGTCTTCCACTG 326

QY 125 TGCCGTCGCGAATGTTCAAGCT 146
Db 327 TTGCTTTGGGAATGAGCAAGGT 348

RESULT 7
US-08-708-123D-1
; Sequence 1, Application US/08708123D
; Patent No. 6482927
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,123D
; FILING DATE: 03-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
```

```

; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
; US-08-708-123D-1

Query Match 15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTGCTCTTTATTGGTTTTACGCCACTATTCGGGATAAACAATATTCGGCT 124
DB 267 AGCTAAATTAAATCAAGTGTATCTACGTTCTCGAGTTATCCAAAACAGTCTCCACTG 326

QY 125 TGCCTGCGGAATGTTCAAGCT 146
DB 327 TTGCTTTGGGAATGACCAAGGT 348

RESULT 8
US-08-583-153A-1
; Sequence 1, Application US/08583153A
; Patent No. 6506877
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING
; TITLE OF INVENTION: OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,153A
; FILING DATE: 28-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906

```

```

; NAME: Meiklejohn, Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/016001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61...2742
; US-08-583-153A-1

Query Match 15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTGCTCTTTATTGGTTTTACGCCACTATTCGGGATAAACAATATTCGGCT 124
DB 267 AGCTAAATTAAATCAAGTGTATCTACGTTCTCGAGTTATCCAAAACAGTCTCCACTG 326

QY 125 TGCCTGCGGAATGTTCAAGCT 146
DB 327 TTGCTTTGGGAATGACCAAGGT 348

RESULT 9
US-08-570-142D-1
; Sequence 1, Application US/08570142D
; Patent No. 6509189
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING
; TITLE OF INVENTION: OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,142D
; FILING DATE: 11-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906

```

```

/ TELEFAX: 200154
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3097 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: Coding Sequence
/ LOCATION: 61...2742
/ US-08-570-142D-1

Query Match      15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTATTGGTTTTTACGCCCACTCATTCGGGATAAACAATATTCGGCCT 124
Db 267 ACCTAAATTTAATTCAGTGGTATCTACGTTCTCGATTATCCAAACAGCTTCCACTG 326

QY 125 TGGCGTCGCGAATGTTCAAGCT 146
Db 327 TTGCTTTGGGAATGAGCAAGGT 348

RESULT 10
US-08-638-524B-1
/ Sequence 1, Application US/08638524B
/ Patent No. 6548269
/ GENERAL INFORMATION:
/ APPLICANT: Tartaglia, Louis A.
/ APPLICANT: Tepper, Robert I.
/ APPLICANT: Culpepper, Janice A.
/ APPLICANT: White, David W.
/ TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR THE
/ TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OB
/ TITLE OF INVENTION: CACHEXIA
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson, P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: US
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows95
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/638,524B
/ FILING DATE: 26-APR-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/599,455
/ FILING DATE: 22-JAN-1996
/ APPLICATION NUMBER: 08/583,153
/ FILING DATE: 28-DEC-1995
/ APPLICATION NUMBER: 08/570,142
/ FILING DATE: 11-DEC-1995
/ APPLICATION NUMBER: 08/569,485
/ FILING DATE: 08-DEC-1995
/ APPLICATION NUMBER: 08/566,622
/ FILING DATE: 04-DEC-1995
/ APPLICATION NUMBER: 08/562,663
/ FILING DATE: 27-NOV-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meiklejohn Ph.D., Anita L.
/ REGISTRATION NUMBER: 35,283
/ REFERENCE/DOCKET NUMBER: 07334/018001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-542-5070

```

```

/ TELEFAX: 617-542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3097 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: Coding Sequence
/ LOCATION: 61...2742
/ US-08-638-524B-1

Query Match      15.2%; Score 29.2; DB 4; Length 3097;
Best Local Similarity 59.8%; Pred. No. 2.1;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTATTGGTTTTTACGCCCACTCATTCGGGATAAACAATATTCGGCCT 124
Db 267 ACCTAAATTTAATTCAGTGGTATCTACGTTCTCGATTATCCAAACAGCTTCCACTG 326

QY 125 TGGCGTCGCGAATGTTCAAGCT 146
Db 327 TTGCTTTGGGAATGAGCAAGGT 348

RESULT 11
US-08-827-962-17
/ Sequence 17, Application US/08827962A
/ Patent No. 6258944
/ GENERAL INFORMATION:
/ APPLICANT: VERCK & CO., INC.
/ TITLE OF INVENTION: OB RECEPTOR ISOFORMS AND NUCLEIC ACIDS
/ TITLE OF INVENTION: ENCODING THEM
/ FILE REFERENCE: 19693
/ CURRENT APPLICATION NUMBER: US/08/827,962A
/ CURRENT FILING DATE: 1997-05-06
/ PRIOR APPLICATION NUMBER: 60/016,899
/ PRIOR FILING DATE: 1996-05-06
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17
/ LENGTH: 3495
/ TYPE: DNA
/ ORGANISM: Rattus No. 6258944vegicus
/ US-08-827-962-17

Query Match      15.2%; Score 29.2; DB 3; Length 3495;
Best Local Similarity 59.8%; Pred. No. 2.2;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTATTGGTTTTTACGCCCACTCATTCGGGATAAACAATATTCGGCCT 124
Db 281 ACCTAAATTTAATTCAGTGGTATCTACGTTCTCGATTATCCAAACAGCTTCCACTG 340

QY 125 TGGCGTCGCGAATGTTCAAGCT 146
Db 341 TTGCTTTGGGAATGAGCAAGGT 362

RESULT 12
US-08-837-635-5
/ Sequence 5, Application US/08837635
/ Patent No. 600798
/ GENERAL INFORMATION:
/ APPLICANT: ROSENBLUM, CHARLES. I.
/ APPLICANT: VAN DER PLOEG, LEONARDUS, H.T.
/ APPLICANT: QURESHI, SAJJAD, A.
/ APPLICANT: CULLY, DORIS, F.
/ APPLICANT: HESS, JOHN W.
/ APPLICANT: TOTA, MICHAEL, R.
/ APPLICANT: CHEN, FANG

```

```

; TITLE OF INVENTION: LEPTIN ASSAY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOANNE M. GIESSEY - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,635
; FILING DATE: 21-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,051
; FILING DATE: 22-APR-1996
; APPLICATION NUMBER: 60/031,002
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GIESSEY, JOANNE M
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19686Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3046
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3650 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; US-08-837-635-5

Query Match 15.2%; Score 29.2; DB 3; Length 3650;
Best Local Similarity 59.8%; Pred. No. 2.3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTTATTTAGCCACCTATTCGGGATTAACCAATATTCGGCT 124
Db 281 AGCTAAATTTAATTCACCTGTTATCTACGTTCTGAGTTATCCAAACCAATTTCCACTG 340

QY 125 TCCGCTCGGAATGTTCAAGCT 146
Db 341 TTGCTTTGGGAATGACCAAGT 362

RESULT 13
US-08-837-635-7
; Sequence 7, Application US/08837635
; Patent No. 6007998
; GENERAL INFORMATION:
; APPLICANT: ROSENBLUM, CHARLES. I.
; APPLICANT: VAN DER PLOEG, LEONARDUS, H.T.
; APPLICANT: QURESHI, SAJJAD, A.
; APPLICANT: CULLY, DORIS, F.
; APPLICANT: HESS, JOHN W.
; APPLICANT: TOTA, MICHAEL, R.
; APPLICANT: CHEN, FANG
; TITLE OF INVENTION: LEPTIN ASSAY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOANNE M. GIESSEY - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA

; TITLE OF INVENTION: LEPTIN ASSAY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOANNE M. GIESSEY - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,635
; FILING DATE: 21-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,051
; FILING DATE: 22-APR-1996
; APPLICATION NUMBER: 60/031,002
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GIESSEY, JOANNE M
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19686Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3046
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3650 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; US-08-837-635-7

Query Match 15.2%; Score 29.2; DB 3; Length 3650;
Best Local Similarity 59.8%; Pred. No. 2.3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTTATTTAGCCACCTATTCGGGATTAACCAATATTCGGCT 124
Db 281 AGCTAAATTTAATTCACCTGTTATCTACGTTCTGAGTTATCCAAACCAATTTCCACTG 340

QY 125 TCCGCTCGGAATGTTCAAGCT 146
Db 341 TTGCTTTGGGAATGACCAAGT 362

RESULT 14
US-08-827-962-16
; Sequence 16, Application US/08827962A
; Patent No. 6258944
; GENERAL INFORMATION:
; APPLICANT: MERCK & CO., INC.
; TITLE OF INVENTION: OB RECEPTOR ISOFORMS AND NUCLEIC ACIDS
; FILE REFERENCE: 19693
; CURRENT APPLICATION NUMBER: US/08/827,962A
; CURRENT FILING DATE: 1997-05-06
; PRIOR APPLICATION NUMBER: 60/016,899
; PRIOR FILING DATE: 1996-05-06
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3650
; TYPE: DNA
; ORGANISM: Rattus No. 6258944vegicus
; US-08-827-962-16

Query Match 15.2%; Score 29.2; DB 3; Length 3650;
Best Local Similarity 59.8%; Pred. No. 2.3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTGCTTTTATTTAGCCACCTATTCGGGATTAACCAATATTCGGCT 124
Db 281 AGCTAAATTTAATTCACCTGTTATCTACGTTCTGAGTTATCCAAACCAATTTCCACTG 340
```

QY 125 TGCCGTGCGGAATGTTCAAGCT 146
 Db 341 TTGCTTTGGGAATGACCAAGT 362

RESULT 15

US-08-803-346-2
 ; Sequence 2, Application US/08803346
 ; Patent No. 6281346
 ; GENERAL INFORMATION:
 ; APPLICANT: HESS, JOHN W.
 ; APPLICANT: CASKEY, C. THOMAS
 ; APPLICANT: LIU, QINGYUN
 ; APPLICANT: PHILLIPS, MICHAEL SEAN
 ; TITLE OF INVENTION: RAT OB RECEPTORS AND NUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THEM
 ; NUMBER OF SEQUENCES: 77
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: JOANNE M. GIESSE - MERCK & CO., INC.
 ; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
 ; CITY: RAHWAY
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: Fast-Seq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/803,346
 ; FILING DATE: 20-FEB-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: GIESSE, JOANNE M.
 ; REGISTRATION NUMBER: 32,838
 ; REFERENCE/DOCKET NUMBER: 19642Y
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-3046
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3650 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-803-346-2

Query Match 15.2%; Score 29.2; DB 3; Length 3650;
 Best Local Similarity 59.8%; Pred. No. 2.3;
 Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 65 ATCTAAATTTCTTTTATTTTTCAGCCACTTCATTCGGATAAACAATATTCGCCT 124
 Db 281 AGCTAAATTTTAACTCACTGATCTCTGTTTCTGATATCCAAACCACTTTTCACTG 340
 QY 125 TGCCGTGCGGAATGTTCAAGCT 146
 Db 341 TTGCTTTGGGAATGACCAAGT 362

Search completed: November 15, 2003, 08:08:52
 Job time : 36.9056 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 108.726 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-67

Perfect score: 192

Sequence: 1 GATCGCGTACTTGGTTT.....CGGTACGACGCTTCGAGA 192

Scoring table:

IDENTITY NUC

Gapop 10_0 , Gapext 1.0

2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US03_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	192	10	US-09-928-457-67
2	162.4	84.6	229	10	Sequence 67, Appl
3	35.4	18.4	12100	11	Sequence 79, Appl
4	35.4	18.4	12100	11	Sequence 9126, Ap
5	32.8	17.1	74586	9	Sequence 702, App
6	32.8	17.1	74586	9	Sequence 702, App
7	31.6	16.5	3891	13	Sequence 3, Appli
8	31	16.1	15400	12	Sequence 348, App
9	30	15.6	197	10	Sequence 3, Appli
10	29.8	15.5	1136	12	Sequence 6709, Ap
11	29.8	15.5	1136	13	Sequence 9875, Ap
12	29.8	15.5	3458	12	Sequence 9875, Ap
13	29.6	15.4	3782	12	Sequence 70, Appl
14	29.6	15.4	3782	12	Sequence 260325,
15	29.6	15.4	3782	13	Sequence 260326,
16	29.6	15.4	3782	13	Sequence 260326,

17	29.4	15.3	364	12	US-09-814-353-6222	Sequence 6322, Ap
18	29.4	15.3	364	12	US-09-814-353-12600	Sequence 12600, A
19	29.4	15.3	560	11	US-09-918-995-33163	Sequence 33163, A
20	29.4	15.3	562	12	US-10-027-633-47595	Sequence 47595, A
21	29.4	15.3	562	13	US-10-027-633-47595	Sequence 47595, A
22	29.4	15.3	625	12	US-10-027-633-68888	Sequence 68888, A
23	29.4	15.3	625	12	US-10-027-633-294706	Sequence 294706,
24	29.4	15.3	625	13	US-10-027-633-68888	Sequence 68888, A
25	29.4	15.3	625	13	US-10-027-633-294706	Sequence 294706,
26	29.2	15.2	261	9	US-09-912-020-183	Sequence 183, App
27	29.2	15.2	729	12	US-10-027-633-19039	Sequence 19039, A
28	29.2	15.2	729	13	US-10-027-633-19039	Sequence 19039, A
29	29.2	15.2	2679	14	US-10-278-733-14	Sequence 14, Appl
30	29.2	15.2	2679	8	US-08-779-457-8	Sequence 8, Appli
31	29.2	15.2	2868	14	US-10-214-802-8	Sequence 8, Appli
32	29.2	15.2	3097	13	US-10-079-623-1	Sequence 1, Appli
33	29.2	15.2	3489	14	US-10-278-733-18	Sequence 18, Appl
34	29.2	15.2	3656	14	US-10-226-579-1	Sequence 1, Appli
35	29.2	15.2	3854	13	US-10-079-625-42	Sequence 42, Appl
36	29.2	15.2	33239	10	US-09-814-950-3	Sequence 3, Appli
37	29.2	15.2	132762	11	US-09-954-558-17	Sequence 17, Appl
38	29	15.1	645	14	US-10-198-846-7605	Sequence 7605, Ap
39	28.8	15.0	999	11	US-09-769-787-202	Sequence 202, App
40	28.8	15.0	4054	10	US-09-070-927A-251	Sequence 251, App
41	28.8	15.0	16602	12	US-10-311-455-700	Sequence 700, App
42	28.6	14.9	107	9	US-09-864-761-18661	Sequence 18661, A
43	28.6	14.9	398	11	US-09-803-719-703	Sequence 703, App
44	28.6	14.9	476	9	US-09-864-761-1912	Sequence 1912, Ap
45	28.6	14.9	476	9	US-09-864-761-5176	Sequence 5176, Ap

ALIGNMENTS

RESULT 1

US-09-928-457-67
; Sequence 67, Application US/09928457
; Patent No. US20020164603A1

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: DNA, specific proteins and peptides
of the Neisseria meningitidis species bacteria, method
for obtaining them and their biological application.

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 99

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (OEB)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/928,457

FILING DATE: 2001-08-14

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/214,759

FILING DATE: 199-12-10

INFORMATION FOR SEQ ID NO: 67:

SEQUENCE CHARACTERISTICS:

LENGTH: 192 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-928-457-67

Query Match 100.0%; Score 192; DB 10; Length 192;

Best Local Similarity 100.0%; Pred. No. 2e-52;

Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGCGTACTTGGTTTTCATATTTTGCATAGTCTTGTGCGGCGCATCTCCCGA 60

DB 1 GATCGCGTACTTGGTTTTCATATTTTGCATAGTCTTGTGCGGCGCATCTCCCGA 60

QY 61 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCGGATTAACAATATTC 120
 Db |||||
 QY 61 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCGGATTAACAATATTC 120
 Db |||||
 QY 121 GCCTTCCGCTCGGAATGTTCAAGCTAGCTGATCAGCTATCAGGTGCGCGTTACC 180
 Db |||||
 QY 121 GCCTTCCGCTCGGAATGTTCAAGCTAGCTGATCAGCTATCAGGTGCGCGTTACC 180
 Db |||||
 QY 181 GAGCCTTCGAGA 192
 Db |||||
 QY 181 GAGCCTTCGAGA 192
 Db |||||

RESULT 2

US-09-928-457-79/c
 ; Sequence 79, Application US/09928457
 ; Patent No. US20020164603A1
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: DNA, specific proteins and peptides
 ; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
 ; TITLE OF INVENTION: for obtaining them and their biological application.
 ; NUMBER OF SEQUENCES: 99
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (OEB)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/928,457
 ; FILING DATE: 2001-08-14
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/214,759
 ; FILING DATE: 199-12-10
 ; INFORMATION FOR SEQ ID NO: 79:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 229 base pairs
 ; TYPE: nucleotide
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-928-457-79

Query Match 84.6%; Score 162.4; DB 10; Length 229;
 Best Local Similarity 93.8%; Pred. No. 9e-43; Indels 1; Gaps 1;
 Matches 180; Conservative 0; Mismatches 11;
 QY 1 GATCCGCTACTTGGTTTTCATATTTTGCATAGTCTTGTGCGTGGGCAATCTTCCCGA 60
 Db |||||
 QY 204 GATCCGCTACTTGGTTTTCATATTTTGCATAGTCTTGTGCGTGGGCAATCTTCCCGA 145
 Db |||||
 QY 61 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCGGATTAACAATATTC 120
 Db |||||
 QY 144 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCGGATTAACAATATTC 86
 Db |||||
 QY 121 GCCTTCCGCTCGGAATGTTCAAGCTAGCTGATCAGCTATCAGGTGCGCGTTACC 180
 Db |||||
 QY 85 GCCTTCCGCTCGGAATGTTCAAGCTAGCTGATCAGCTATCAGGTGCGCGTTACC 26
 Db |||||
 QY 181 GAGCCTTCGAGA 192
 Db |||||
 QY 25 GAGCCTTCGAGA 14
 Db |||||

RESULT 3

US-09-764-891-9126/c
 ; Sequence 9126, Application US/09764891
 ; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC006
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 9126
 ; LENGTH: 12100
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-891-9126

Query Match 18.4%; Score 35.4; DB 11; Length 12100;
 Best Local Similarity 51.6%; Pred. No. 0.84;
 Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 1 GATCCGCTACTTGGTTTTCATATTTTGCATAGTCTTGTGCGTGGGCAATCTTCCCGA 60
 Db 4343 GCTCGTAGTACTGGTTGTTATTTTCTCTGGGTCAGTGGTAACATTCCTT 4284
 QY 61 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCGGATTAACAATATTC 120
 Db 4283 CATCATCTAAATTTGCTTTTATTTGTTTATGCTTCTGCTCATCTTCTTTATTAATCTAGTACT 4224
 QY 121 GCCTTCCGCTCGGAATGTTCAAGCTAGCTAGCTGCATCA 157
 Db 4223 GCCTATCTTAGTAATTTTTCAAAAACCTTGATTA 4187

RESULT 4

US-10-091-572-702/c
 ; Sequence 702, Application US/10091572
 ; Publication No. US20030054373A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PALL8C1
 ; CURRENT APPLICATION NUMBER: US/10/091,572
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 09/764,850
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION NUMBER: 60/179,065
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: 60/180,628
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: 60/214,886
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 60/217,487
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/225,758
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/220,963
 ; PRIOR FILING DATE: 2000-07-26
 ; PRIOR APPLICATION NUMBER: 60/217,496
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/225,447
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/218,290
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/225,757
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/226,868
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 60/216,647
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/225,267
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/216,880
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/225,270
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/251,869

PRIOR FILING DATE: 2000-12-08	
PRIOR APPLICATION NUMBER: 60/235,834	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: 60/234,274	
PRIOR FILING DATE: 2000-09-21	
PRIOR APPLICATION NUMBER: 60/234,223	
PRIOR FILING DATE: 2000-09-21	
PRIOR APPLICATION NUMBER: 60/228,924	
PRIOR FILING DATE: 2000-08-30	
PRIOR APPLICATION NUMBER: 60/224,518	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/236,369	
PRIOR FILING DATE: 2000-09-29	
PRIOR APPLICATION NUMBER: 60/224,519	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/220,964	
PRIOR FILING DATE: 2000-07-26	
PRIOR APPLICATION NUMBER: 60/241,809	
PRIOR FILING DATE: 2000-10-20	
PRIOR APPLICATION NUMBER: 60/249,299	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/236,327	
PRIOR FILING DATE: 2000-09-29	
PRIOR APPLICATION NUMBER: 60/241,785	
PRIOR FILING DATE: 2000-10-20	
PRIOR APPLICATION NUMBER: 60/244,617	
PRIOR FILING DATE: 2000-11-01	
PRIOR APPLICATION NUMBER: 60/225,268	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/236,368	
PRIOR FILING DATE: 2000-09-29	
PRIOR APPLICATION NUMBER: 60/251,856	
PRIOR FILING DATE: 2000-12-08	
PRIOR APPLICATION NUMBER: 60/251,868	
PRIOR FILING DATE: 2000-12-08	
PRIOR APPLICATION NUMBER: 60/229,344	
PRIOR FILING DATE: 2000-09-01	
PRIOR APPLICATION NUMBER: 60/234,997	
PRIOR FILING DATE: 2000-09-25	
PRIOR APPLICATION NUMBER: 60/229,343	
PRIOR FILING DATE: 2000-09-01	
PRIOR APPLICATION NUMBER: 60/229,345	
PRIOR FILING DATE: 2000-09-01	
PRIOR APPLICATION NUMBER: 60/229,287	
PRIOR FILING DATE: 2000-09-01	
PRIOR APPLICATION NUMBER: 60/229,513	
PRIOR FILING DATE: 2000-09-05	
PRIOR APPLICATION NUMBER: 60/231,413	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/229,509	
PRIOR FILING DATE: 2000-09-05	
PRIOR APPLICATION NUMBER: 60/236,367	
PRIOR FILING DATE: 2000-09-29	
PRIOR APPLICATION NUMBER: 60/237,039	
PRIOR FILING DATE: 2000-10-02	
PRIOR APPLICATION NUMBER: 60/237,038	
PRIOR FILING DATE: 2000-10-02	
PRIOR APPLICATION NUMBER: 60/236,370	
PRIOR FILING DATE: 2000-09-29	
PRIOR APPLICATION NUMBER: 60/236,802	
PRIOR FILING DATE: 2000-10-02	
PRIOR APPLICATION NUMBER: 60/237,037	
PRIOR FILING DATE: 2000-10-02	
PRIOR APPLICATION NUMBER: 60/237,040	
PRIOR FILING DATE: 2000-10-02	
PRIOR APPLICATION NUMBER: 60/240,960	
PRIOR FILING DATE: 2000-10-20	
PRIOR APPLICATION NUMBER: 60/239,935	
PRIOR FILING DATE: 2000-10-13	
PRIOR APPLICATION NUMBER: 60/239,937	
PRIOR FILING DATE: 2000-10-13	
PRIOR APPLICATION NUMBER: 60/241,787	
PRIOR FILING DATE: 2000-10-20	

PRIOR APPLICATION NUMBER: 60/246,474	PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532	PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532	PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681	PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759	PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213	PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182	PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214	PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,833	PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438	PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135	PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266	PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297	PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400	PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242	PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081	PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080	PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414	PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244	PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064	PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063	PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397	PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399	PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401	PRIOR FILING DATE: 2000-09-14

```
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      18.4%; Score 35.4; DB 14; Length 12100;
Best Local Similarity 51.6%; Pred. No. 0.84;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 GATCCGCTACTTGGTTTTCATATTTCATAGTCTTGTGCGTGGCGCATCTTCCCGA 60
Db 4343 GCTCGTAGTACTGTTGTTATTTTCTCTGGGTCACTGTAACATTCCTT 4284
QY 61 CATCATCTAAATTCCTTTTATTTGTTTTCAGCCACATCTTGGCGATAAACAATATTC 120
Db 4283 CATCATCTAAATTCCTTTTATTTGAACTCTCTCATCTCTTTATTAATCTAGTAGT 4224
QY 121 GCCTTCCGTCGCGAATGTTCAAGCTAGCTGCATCA 157
Db 4223 GCCTATCTTAGTAATTTTCAAAAACCTTGATTA 4187

RESULT 5
US-09-781-558-3/c
; Sequence 3, Application US/09781558
; Patent No. US20020034778A1
; GENERAL INFORMATION:
; APPLICANT: Beasley et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000756
; CURRENT APPLICATION NUMBER: US/09/781,558
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/199,811
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/641,426
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 74586
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(74586)
; OTHER INFORMATION: n = A,T,C or G
US-09-781-558-3

Query Match      17.1%; Score 32.8; DB 9; Length 74586;
Best Local Similarity 53.0%; Pred. No. 12;
Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 15 GTTTTCATATTTTCATAGTCTTGTGCGTGGCGATCTTCCCGACATCATCTAAATTT 74
Db 48079 GTACTTCATCTGTCTAAACTTTGGCGTATTGTAACCTTCTAAGTTGCTTCTCTTTT 48020
QY 75 GCTTTTATGTTTTCAGCCACTCATTCGGGATAAACAATATTCGCTTCCGCTCGCG 134
Db 48019 GTTTATGCTTTTCTCCCTAATTCATTGAAGTCAACAATTTGACTCATTTACCTCTGA 47960
QY 135 AATGTTCAAGCT 146
Db 47959 ATTGCTCTACCT 47948

RESULT 7
US-10-108-605-348/c
; Sequence 348, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 348
; LENGTH: 3891
; TYPE: DNA
```

```
Db 47959 ATTGCTCTACCT 47948

RESULT 6
US-10-388-650-3/c
; Sequence 3, Application US/10388650
; Publication No. US20030166151A1
; GENERAL INFORMATION:
; APPLICANT: BRANDON et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000756-CIPCON
; CURRENT APPLICATION NUMBER: US/10/388,650
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/199,811
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/641,426
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 09/781,558
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 74586
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(74586)
; OTHER INFORMATION: n = A,T,C or G
US-10-388-650-3

Query Match      17.1%; Score 32.8; DB 12; Length 74586;
Best Local Similarity 53.0%; Pred. No. 12;
Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 15 GTTTTCATATTTTCATAGTCTTGTGCGTGGCGATCTTCCCGACATCATCTAAATTT 74
Db 48079 GTACTTCATCTGTCTAAACTTTGGCGTATTGTAACCTTCTAAGTTGCTTCTCTTTT 48020
QY 75 GCTTTTATGTTTTCAGCCACTCATTCGGGATAAACAATATTCGCTTCCGCTCGCG 134
Db 48019 GTTTATGCTTTTCTCCCTAATTCATTGAAGTCAACAATTTGACTCATTTACCTCTGA 47960
QY 135 AATGTTCAAGCT 146
Db 47959 ATTGCTCTACCT 47948

RESULT 7
US-10-108-605-348/c
; Sequence 348, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 348
; LENGTH: 3891
; TYPE: DNA
```

```
; ORGANISM: Drosophila melanogaster
US-10-108-605-348

Query Match      16.5%; Score 31.6; DB 13; Length 3891;
Best Local Similarity 52.2%; Pred. No. 9.1;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 9 TACTTGGTTTTCATATTTTGATAGTCTTGTGCGTGGGCATCTTCCCGACATCATCT 68
Db 1290 TTCTCGCGTGTGCTGTAATCACTTGTCTGTGCTGTCTGCGAGTTTTCGCCATCTCGGT 1231

QY 69 AAATTGCTTTTATTTGGTTTTTACGCCACTATTCGGGATAAACAATTCGGCCTTCCC 128
Db 1230 AGACATGGCTTGAGCGGCGCTTCCTCCAGTCATTCGTCATCCAGATCGGTCAATGTC 1171

QY 129 GTCGGAATGTTCA 142
Db 1170 ACCTTAATGGCA 1157

RESULT 8
US-10-090-002-3/c
; Sequence 3, Application US/10090002
; Publication No. US20030175926A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001316
; CURRENT APPLICATION NUMBER: US/10/090,002
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-090-002-3.

Query Match      16.1%; Score 31; DB 12; Length 15400;
Best Local Similarity 55.0%; Pred. No. 25;
Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 81 ATTGTTTTAGCCACATTCGCGATAAACAATATTCGCGCTTGGCGTTCGCGAATGTT 140
Db 2589 AGTGATGCCCGCCGCCACCCCTGGGATGTGTGGGGTGTCTCCCTTCCCTCTCTCC 2530

QY 141 CAAGTAGCGCTGCATCACCGTAATCAGGTGCGCGTTACCGAGCCTTCGAG 191
Db 2529 CCGACTCCCGCGCTCTCGAATCAGGTTTCCAGCAGCTGGGCTTCAG 2479

RESULT 9
US-09-974-300-6709/c
; Sequence 6709, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6709
; LENGTH: 197

; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6709

Query Match      15.8%; Score 30; DB 10; Length 197;
Best Local Similarity 61.8%; Pred. No. 9;
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 38 TGTCGGTGGGCATCTTCCCGACATCATCTAAATTTGTTTATTGGTTTTTACGCCAC 97
Db 82 TTCTTCCGTTCTTTTGTTCGGCTTAATGTAATTGGCTTTTCTTTTACGCAAC 23

QY 98 TCATTGGGATAAACAAT 115
Db 22 CAATTGCAGTCGCAGAT 5

RESULT 10
US-10-027-632-9875
; Sequence 9875, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9875
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-9875

Query Match      15.8%; Score 29.8; DB 12; Length 1136;
Best Local Similarity 50.3%; Pred. No. 21;
Matches 73; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 3 TCCGCGTACTTGTTTTTCATATTTTGCATAGTCTGTGTCGGTCGGGATCTTCCCGACA 62
Db 874 TCCAGGTTCTTGTTCCTCATTCGGGAACATGGTCTTGTGACGAGATAATCTTGAACCCG 933

QY 63 TCATCTAAATTTGCTTTTATTTGTTTTTACGCCACATTCGCGATAAACAATATTCGCG 122
Db 934 TTCTTAAATATTTCTTTTGGTCAATTTAAAGCCAGCGGAGGGGTGGAAGGCCCTT 993

QY 123 CTTCGCGTCGGAATGTTCAAGCTA 147
Db 994 TATTGCTCAGGAAGTGCATGTA 1018

RESULT 11
US-10-027-632-9875
; Sequence 9875, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 9875
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-9875

```

	Query Match	15.5%	Score 29.8;	DB 13;	Length 1136;
	Best Local Similarity	50.3%;	Pred. No. 21;		
	Matches	73;	Conservative 0;	Mismatches 72;	Indels 0;
	Gaps	0;			
QY	3	TCGGGCTACTGGTTTTCATATTTCATAGTCTTGTGGTCGGGCATCTTCCCGGACA	62		
Db	874	TCAGGTTCTTGTTCCTCATTTGGAAACATGGTCTTGTGGACGGATAATCTTTGAACCCCG	933		
QY	63	TCATCTAAATTGTCTTTTATTGGTTTTTACGCCACTCATTTGGCGATAAACCAATATTCGC	122		
Db	934	TTTCTTAAATATCTTTGGTTCATTTAAAGCCAGCGGAGGGTGGAAAAAGCCCTT	993		
QY	123	CTTGGCTCGCGAATGTTCAGCTA	147		
Db	994	TATTTGCTCAGGAAGGTGCATGTTA	1018		

RESULT 12
US-10-277-161-70
; Sequence 70, Application US/10277161
; Publication No. US20030194696A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Smith, Ernest S.
; TITLE OF INVENTION: Methods of Producing a Library and Methods of Selecting Polynucle
; TITLE OF INVENTION: of Interest
; FILE REFERENCE: 1821.0050006
; CURRENT APPLICATION NUMBER: US/10/277,161
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/192,586
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/203,343
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 60/263,226
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/271,426
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/818,991
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 70
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-277-161-70

Query Match	15.5%	Score 29.8	DB 12	Length 3458
Best Local Similarity	50.3%	Pred. No. 33		
Matches	73	Conservative 0	Mismatches 72	Indels 0
Gaps	0			
Qy	3	TCGCGTACTCGTGTTCATATTTTGCATAGTCTGTGGTGGGCATCTTCCCCGACA	62	
Db	1834	TCAGGTTCTGTTCCTCATTTGGGAACATGGTCITTTGGACGGATAATCTTGAACCCCG	1893	
Qy	63	TCATCTAAATTTGCTCTTTATTTGGTTTTTACGCCATCATTTGGGATAAACAATATTTCCGC	122	
Db	1894	TITCTAAATATTCCTTGGTCATTTAAGCCACGCGGAGGGTGGAAAAAGCCCT	1953	
Qy	123	CTTGGCGTCGCGAATGTTCAAGCTA	147	
Db	1954	TATTTGCTCAGGAAGGTGATGTTA	1978	

```

RESULT 13
US-10-027-632-260325
, Sequence 260325, Application US/10027632
, Publication No. US20030204075A9
, GENERAL INFORMATION:
, APPLICANT: Wang, David G.
, TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
, TITLE OF INVENTION: Polymorphisms in the Human Genome
, FILE REFERENCE: 10827.129
, CURRENT APPLICATION NUMBER: US/10/027,632
, CURRENT FILING DATE: 2002-04-30
, PRIOR APPLICATION NUMBER: US 60/218,006
, PRIOR FILING DATE: 2000-07-12
, PRIOR APPLICATION NUMBER: US 60/198,676
, PRIOR FILING DATE: 2000-04-20
, PRIOR APPLICATION NUMBER: US 60/193,483
, PRIOR FILING DATE: 2000-03-29
, PRIOR APPLICATION NUMBER: US 60/185,218
, PRIOR FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/157,363
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: US 60/156,358
, PRIOR FILING DATE: 1999-09-28
, PRIOR APPLICATION NUMBER: US 60/146,002
, PRIOR FILING DATE: 1999-08-09
, NUMBER OF SEQ ID NOS: 325720
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 260325
, LENGTH: 3782
, TYPE: DNA
, ORGANISM: Human
US-10-027-632-260325

```

	Query Match	15.4%;	Score 29.6;	DB 12;	Length 3782;
	Best Local Similarity	50.7%;	Pred. No. 40;		
	Matches	71;	Conservative 0;	Mismatches 69;	Indels 0; Gaps 0;
QY	11	CTTGGCTTTTCATATTTGCATAGCTCTGTGCGTGGGCATCTTCCCGACATCATCTAA	70		
Db	2273	CTCTCTTAATCATCTCTCCCTCCCTGCCACCTCAGTCACCTCTGTTCCCTTATCTGTG	2332		
QY	71	ATTTCGCTTTATGGTTTTTACGCCCATCTATTGCGGATAACAATATTCGGCTTGCCTT	130		
Db	2333	TTTTGTTTTTATCAGATATTCACCTCTTTTTCGGTGAAATATTATTATTATTTGCTGTT	2392		
QY	131	CGCGAATGTTCAAGCTAGCC	150		
Db	2393	TTCTAAATGTCAGTCTTGCC	2412		

RESULT 14
US-10-027-632-260326
; Sequence 260326; Application US/10027632
; Publication NO. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260326
; LENGTH: 3782
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-260326

Query Match      15.4%; Score 29.6; DB 12; Length 3782;
Best Local Similarity 50.7%; Pred. No. 40;
Matches 71; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      11  CTTGGTTTTTCATATTTTCATAGTCTTGTGGTGGGCATCTTCCCGACATCACTAA 70
Db      2273  CTTCTTTAATCATCTTCTCCCTCCCTGCCACCTCAGTCACCCCTGTGTCCTTTATCTTG 2332

QY      71  ATTGTCTTTATTGGTTTTACGCCACTCATTGGGATAAACAATATTCGCCCTTGCCGT 130
Db      2333  TTTTGGTTTTATCACAGTATTCACCTCTTTTGGCGTGAATATTATTATTGTCGTGT 2392

QY      131  CGCGAATGTTCAAGCTAGCC 150
Db      2393  TTCTAATTGTCAGTCTTGCC 2412

Search completed: November 15, 2003, 08:31:48
Job time : 110.826 secs.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260325
; LENGTH: 3782
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-260325

Query Match      15.4%; Score 29.6; DB 12; Length 3782;
Best Local Similarity 50.7%; Pred. No. 40;
Matches 71; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      11  CTTGGTTTTTCATATTTTCATAGTCTTGTGGTGGGCATCTTCCCGACATCACTAA 70
Db      2273  CTTCTTTAATCATCTTCTCCCTCCCTGCCACCTCAGTCACCCCTGTGTCCTTTATCTTG 2332

QY      71  ATTGTCTTTATTGGTTTTACGCCACTCATTGGGATAAACAATATTCGCCCTTGCCGT 130
Db      2333  TTTTGGTTTTATCACAGTATTCACCTCTTTTGGCGTGAATATTATTATTGTCGTGT 2392

QY      131  CGCGAATGTTCAAGCTAGCC 150
Db      2393  TTCTAATTGTCAGTCTTGCC 2412

RESULT 15
US-10-027-632-260325
; Sequence 260325, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260325
; LENGTH: 3782
; TYPE: DNA
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 531.201 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-74
Perfect score: 120
Sequence: 1 CGGTGAGAACAGGCAAGCT.....CTCAAGGAGAGAGCTATCAT 120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.in.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ov.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.in.*
32: em.htg.inv.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vit.*
38: em.sy.*
39: em.hgo.hum.*
40: em.hgo.mus.*
41: em.hgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	120	100.0	120	6	A68902	A68902 Sequence 73
2	120	100.0	120	6	A68903	A68903 Sequence 74
3	120	100.0	120	6	BD063002	BD063002 DNA and s
4	120	100.0	120	6	BD063003	BD063003 DNA and s
5	55.8	46.5	354	1	AF169447	AF169447 Neisseria
6	55.8	46.5	2094	6	AX024108	AX024108 Sequence
7	55.8	46.5	340806	1	NNA122491	AX024133 Neisseria
8	55	45.8	2094	6	AX024133	AX024133 Sequence
9	55	45.8	10467	1	AE002381	AE002381 Neisseria
10	55	45.8	349980	6	AX043922	AX043922 Sequence
11	32.8	27.3	223203	2	AC113844	AC113844 Rattus no
12	31.8	26.5	174633	2	EX537269	EX537269 Dario rer
13	30.2	25.2	187986	2	AC138366	AC138366 Mus muscu
14	30.2	25.2	254098	2	AC096226	AC096226 Rattus no
15	30	25.0	204173	2	AC137237	AC137237 Rattus no
16	30	25.0	207918	2	AC112051	AC112051 Rattus no
17	30	25.0	258286	2	AC094841	AC094841 Rattus no
18	29.8	24.8	227385	2	AC094407	AC094407 Rattus no
19	29.8	24.8	235915	2	AC111641	AC111641 Rattus no
20	29.8	24.8	238377	2	AC125923	AC125923 Rattus no
21	29.8	24.8	240497	2	EX072570	EX072570 Dario rer
22	29.8	24.8	251180	2	AC123665	AC123665 Rattus no
23	29.8	24.8	279615	2	AC103332	AC103332 Rattus no
24	29.6	24.7	2635	8	AY093366	AY093366 Arabidops
25	29.6	24.7	2987	8	AY072210	AY072210 Arabidops
26	29.6	24.7	12519	1	AE006089	AE006089 Pasteurel
27	29.6	24.7	99254	8	AC002423	AC002423 Genomic s
28	29.6	24.7	199001	2	AC136115	AC136115 Rattus no
29	29.6	24.7	209574	2	AC102170	AC102170 Mus muscu
30	29.4	24.5	22018	2	AC125975	AC125975 Rattus no
31	29.4	24.5	230956	2	AC105782	AC105782 Rattus no
32	29.4	24.5	236290	2	AC114863	AC114863 Rattus no
33	29.4	24.5	239575	2	AC093194	AC093194 Rattus no
34	29.4	24.5	274620	2	AC126199	AC126199 Rattus no
35	29.2	24.3	10310	1	AE008086	AE008086 Agrobacte
36	29.2	24.3	10947	1	AE009121	AE009121 Agrobacte
37	29.2	24.3	140136	9	AL136087	AL136087 Human DNA
38	29.2	24.3	142711	10	AC121863	AC121863 Mus muscu
39	29.2	24.3	179668	10	AC124399	AC124399 Mus muscu
40	29.2	24.3	187191	9	AC007222	AC007222 Homo sapi
41	29.2	24.3	196181	2	AC144665	AC144665 Rattus no
42	29.2	24.3	221146	10	AC098739	AC098739 Mus muscu
43	29.2	24.3	237569	2	AC130141	AC130141 Rattus no
44	29.2	24.3	235653	2	AC110482	AC110482 Rattus no
45	29.2	24.3	237985	2	AC125754	AC125754 Rattus no

ALIGNMENTS

RESULT 1
A68902 LOCUS A68902 Sequence 73 from Patent WO9802547.
DEFINITION A68902 linear DNA 120 bp PAT 06-MAY-1999
ACCESSION A68902
VERSION A68902.1 GI:4759821
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 120)
AUTHORS Nassif, X., Tinsley, C., Achtman, M., Ruelle, J., Vinals, C. and
Marker, P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL

APPLICATIONS
 Patent: WO 9802547-A 73 22-JAN-1998;
 INST NAT SANTE RECH MED (FR)
 Other publication FR 2751000 19980116.
 Location/Qualifiers
 1..120
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644" 22 t
 38 a 30 c 30 g

JOURNAL
 Patent: WO 9802547-A 73 22-JAN-1998;
 INST NAT SANTE RECH MED (FR)
 Other publication FR 2751000 19980116.
 Location/Qualifiers
 1..120
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644" 22 t
 38 a 30 c 30 g

BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 120; DB 6; Length 120;
 Best Local Similarity 100.0%; Pred. No. 6.3e-31;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 DB 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 QY 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120
 DB 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120

RESULT 2
 A68903 120 bp DNA linear PAT 06-MAY-1999
 LOCUS
 DEFINITION Sequence 74 from Patent WO9802547.
 A68903
 ACCESSION
 VERSION A68903.1 GI:4759822
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 120)
 AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and Merker,P.
 TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL APPLICATIONS
 JOURNAL Patent: WO 9802547-A 74 22-JAN-1998;
 INST NAT SANTE RECH MED (FR)
 Other publication FR 2751000 19980116.
 Location/Qualifiers
 1..120
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644" 22 t
 38 a 30 c 30 g

BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 120; DB 6; Length 120;
 Best Local Similarity 100.0%; Pred. No. 6.3e-31;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 DB 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 QY 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120
 DB 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120

RESULT 3
 BD063002 120 bp DNA linear PAT 27-AUG-2002
 LOCUS
 DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.
 BD063002
 ACCESSION
 VERSION BD063002.1 GI:22608605

KEYWORDS JP 2001504684-A/64.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 120)
 AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.
 TITLE DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications
 JOURNAL Patent: JP 2001504684-A 64 10-APR-2001;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM
 COMMENT PN JP 2001504684-A/64
 PD 10-APR-2001
 PF 11-JUL-1997 JP 1998505685
 PR 12-JUL-1996 FR 96/08768
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI CARLA VINALS.
 PI PETRA MERKER
 PC C12N15/31.C07K14/22.C07K16/12.A61K39/095.C12Q1/68.G01N33/53 CC
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 source
 1..120
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644" 22 t
 38 a 30 c 30 g

BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 120; DB 6; Length 120;
 Best Local Similarity 100.0%; Pred. No. 6.3e-31;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 DB 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTCGGAACGAAA 60
 QY 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120
 DB 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGGCTATCAT 120

RESULT 4
 BD063003 120 bp DNA linear PAT 27-AUG-2002
 LOCUS
 DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.
 BD063003
 ACCESSION
 VERSION BD063003.1 GI:22608606
 KEYWORDS JP 2001504684-A/65.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 120)
 AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.
 TITLE DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications
 JOURNAL Patent: JP 2001504684-A 65 10-APR-2001;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM
 COMMENT PN JP 2001504684-A/65
 PD 10-APR-2001
 PF 11-JUL-1997 JP 1998505685
 PR 12-JUL-1996 FR 96/08768
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI

```

CARLA VINALS,
PI PETRA MERKER
PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
    source
        1..120
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
BASE COUNT      38 a      30 c      30 g      22 t
ORIGIN
Query Match      100.0%; Score 120; DB 6; Length 120;
Best Local Similarity 100.0%; Pred. No: 6.3e-31;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAA 60
    |||
Db 1 CGGTGAGAACAGGCAAGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAA 60
    |||

QY 61 ACTCCTTACCGAAGTCTTCTATATACCCAGGCTCAATAGCGGCTCAAGGAGAGAGCTATCAT 120
    |||
Db 61 ACTCCTTACCGAAGTCTTCTATATACCCAGGCTCAATAGCGGCTCAAGGAGAGAGCTATCAT 120
    |||

RESULT 5
AF169447
LOCUS      AF169447      354 bp      DNA      linear      BCT 09-AUG-2000
DEFINITION      Neisseria meningitidis strain Z2491 clone Cn020 unknown sequence.
ACCESSION      AF169447
VERSION      AF169447.1 GI:9754655
KEYWORDS
SOURCE
    ORGANISM
        Neisseria meningitidis
        Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
        Neisseriaceae; Neisseria.
REFERENCE
    1 (bases 1 to 354)
    Perrin A., Nasif X. and Tinsley C.R.
    Identification of regions of the chromosome of Neisseria
    meningitidis and Neisseria gonorrhoeae which are specific to
    pathogenic Neisseriae
    Unpublished
JOURNAL
REFERENCE
    2 (bases 1 to 354)
    Perrin A., Nasif X. and Tinsley C.R.
    Direct Submission
    Submitted (16-JUN-1999) Necker Medicine Faculty, INSERM U411, 156
    rue de Vaugirard, Paris 75015, France
    Location/Qualifiers
FEATURES
    source
        1..354
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /strain="Z2491"
            /db_xref="taxon:487"
            /clone="Cn020"
BASE COUNT      122 a      77 c      81 g      73 t      1 others
ORIGIN
Query Match      46.5%; Score 55.8; DB 1; Length 354;
Best Local Similarity 79.8%; Pred. No. 2.2e-08;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 17 AGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTACCG---AA 73
    |||
Db 20 AGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTACCGCAA 79
    |||

QY 74 GTCTTCTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 112
    |||
Db 80 GACTTCTATACCCAAAGCTCAATAGCGGCTCAAGGAGAGA 118
    |||

RESULT 6
AF169447
LOCUS      AF169447      2094 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION      Sequence 51 from Patent FR2785293.
ACCESSION      AX024108
VERSION      AX024108.1 GI:10184420
KEYWORDS
SOURCE
    ORGANISM
        Neisseria meningitidis
        Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
        Neisseriaceae; Neisseria.
REFERENCE
    1
    Nassif X., Tinsley C., Aujame L., Perrin A., Rokbi B.,
    Bouchardon A. and Renaud M.G.
    Patent: FR 2785293-A 51 05-MAY-2000;
    PASTEUR MERIEUX SERUMS VACC (FR)
    Location/Qualifiers
FEATURES
    source
        1..2094
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /db_xref="taxon:487"
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=1
            /protein_id="CAC08930.1"
            /db_xref="GI:10184421"
            /translation="MNSTASKTLKGLSLVFFASGFCALIQVSWQRLLFESHGIDLSS
            ITVIISFVMVGLGVGAYFGRIADRFPSIIPLFCIAEVSIGLGLVSRGLISGLGHL
            LVEADLPIIAANFILLPLTFMNGATLPLTFCFNKHNHNGESIGTLYFFNTLGA
            LGLAAAEFFVFFLTQITALTACFNLLIAASVCCVTERMDIVNTKPTSLYMLSF
            LSGLLSGLEVLWVRMFSQAQVPOAFTFLAYFLTGIAGVAYFGKRCRFRVDIP
            FIGCFLWAGIADFLILGAALITLGSFVHAGIFITLSAVVUGLIPLVHHVGTG
            NKSROVSNVYFANVAGSLGPVLIGLDELSTQIILLICLIANVPLPCTLFQK
            SLKNAVSAVSLMFGILMFLPDPSVFQMIADRPDLIENKGIIVAYHRDQKVYIG
            ANVYDGAINTDFNSVNGIERAYLLPSLKGIRIFVWGLSTGSRWARVLSAIPQMOSM
            IVAENPAYSRTADEPQAPLLQDXRVEIVLDGDKWLRHDEDFDILMNTWY
            RAYTNLLSAEFLKQVSHLTDPGIVMFNTHSPHAFATAVHSIPVARYGHMVGSA
            TPVFPNKELLKORSLWIPESGRHVFDSITVDAAAQKVSRMLQIOWTEPSAGAEVI
            TDDNMLVEYKVRGI"
BASE COUNT      428 a      500 c      512 g      654 t
ORIGIN
Query Match      46.5%; Score 55.8; DB 6; Length 2094;
Best Local Similarity 79.8%; Pred. No. 2e-08;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 17 AGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTACCG---AA 73
    |||
Db 747 AGTAAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTACCGCAA 688
    |||

QY 74 GTCTTCTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 112
    |||
Db 687 GACTTCTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 649
    |||

RESULT 7
NMA122491
LOCUS      NMA122491      340806 bp      DNA      linear      BCT 02-SEP-2002
DEFINITION      Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 1/7
VERSION      AL162752.2 GI:7378778
KEYWORDS
SOURCE
    ORGANISM
        Neisseria meningitidis Z2491
        Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
        Neisseriaceae; Neisseria.
REFERENCE
    1 (bases 1 to 340806)
    Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
    Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
    Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N.,
    Holroyd S., Jagels K., Leather S., Moule S., Mungall K.,
    Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M.,

```

Skelton, J., Whitehead, S., Spratt, B.G. and Barrell, B.G.
 Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491
 JOURNAL NATURE 404 (6777), 502-506 (2000)
 MEDLINE 20222556
 PUBMED 10761919
 REFERENCE 2 (bases 1 to 340806)
 AUTHORS Parkhill, J.
 TITLE Direct Submission
 JOURNAL Submitted (30-Mar-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk

Notes:

Details of *N. meningitidis* sequencing at the Sanger Centre are available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/N_meningitidis/).

FEATURES

source
 1..340806
 /organism="Neisseria meningitidis Z2491"
 /mol_type="genomic DNA"
 /strain="Z2491"
 /db_xref="taxon:122587"
 /note="serogroup: A"
 complement(38..358)
 /gene="NMA0001"
 complement(38..358)
 /gene="NMA0001"
 /note="NMA0001, unknown, len: 106 aa"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein NMA0001"
 /protein_id="CAB83321.1"
 /db_xref="GI:7378779"
 /db_xref="SPTREMBL:Q9JX93"
 /translation="MSNTQTIRSDFTTSRAWGLDIANNMGTTVRLHWTDQPKYKHV
 NDGEVFVMDGEVDMHYRENGEHEHIVRLKSGDIFYAGIGCTEHVAHPRGRTRILVIEK
 EGSV"
 complement(414..2438)
 /gene="nuoL"
 complement(414..2438)
 /gene="nuoL"
 /EC_number="1.6.5.3"
 /note="NMA0002, nuoL, NADH dehydrogenase I chain L, len: 674 aa; similar to many e.g. NUOL_RHOCA NADH dehydrogenase I chain L (EC 1.6.5.3) (700 aa). Contains Pfam match to entry PF00361 oxidored q1, NADH-Ubiquinone/plastoquinone (complex I), various chains, and to entry PF00662 oxidored q1_N, NADH-Ubiquinone oxidoreductase (complex I), chain 5 N-terminus"
 /codon_start=1
 /transl_table=11
 /product="NADH dehydrogenase I chain L"
 /protein_id="CAB83322.1"
 /db_xref="GI:7378780"
 /db_xref="SWISS-PROT:Q9JX92"
 /translation="MNDTLYLIILVPLAGSLIAGFGNKGIRGAHTHTIIGLAVS
 AVLGAFLWGLNGSRKAFDENVTYLTMTGDLDFSGVLTDTMTAMVVVTGVSVMV
 HYTGYNHDKVQVQRFYSISLFTSMLMLINSNNFIQLFGWEAVGLVSLIGF
 YFKRPSATFNALRNVGDFGLIG:VLAVFGSLRYQDVPAVLNVQNTATQ
 LFGVEMSLITVTLCLLFPVGAMGSAOPPLHVLDPDSMEGPTPLSALIHAATVTAGL
 FVWSMSIYEMSSTALSIVNIGAITALEMGLFCVQNDIKRVVAISTLSQLGYMTV
 ALGASAYAVMNVHTAFKALLFLAAGSAIIIGMHHDQDMHNGNLUKYMPTIWLTV
 LIGLSLTGTFPGFYSGYKSIIEAKYSTLPGSGFAYFAVLASVFTAFYARQFPM
 VFHGEKRWSPPEHSDGHEHGLKNDNPESPLVTLPLILLAVPVSIVIGYAI
 EPMLYGDFFKQDVFVNAADHFTMHMKEEFHGALAMVSHSLHSPVLALAGVLSAWL
 LVYKLPHLPKIAQAFREPVYLVFNKYVLDALYENFAKGTALGTFFKVGDTAID
 NGVNGSARLVCAVAQAQRKVQTGTFITYTAAAVFGVILLGNTFFWGLER"
 1035..1044
 /note="Core DNA uptake sequence: gccgtctgaa"
 /label=DUS
 complement(1116..2030)
 /gene="nuoL"
 /note="Pfam match to entry PF00361 oxidored q1,

gene

CDS

misc_feature
 /transl_table=11
 /product="hypothetical protein NMA0004"
 /protein_id="CAB83324.1"
 /db_xref="GI:7378782"
 /db_xref="SPTREMBL:Q9JQR9"
 /translation="MPSENPIKTKWKSIDESQLHNARLFPESGDIDRIEVTAGLQQ
 ITRYLFGGLYDPAGIREDNISKGGFRFANAYLKEALVKIEQMPERFETIIAKYVE
 MNIAHPFLGEGNRSTRIWLDVLKKNLKVNNQVNSKTYLQAMERSFVNDLELRFLL
 LKNDLITDDVNRREIIFKGIQESYVYEGYK"
 complement(3516..3525)
 /gene="NMA0004"
 /note="Core DNA uptake sequence: gccgtctgaa"
 /label=DUS
 complement(3561..3866)
 /gene="nuoK"
 complement(3561..3866)
 /gene="nuoK"
 /EC_number="1.6.5.3"
 /note="NMA0005, nuoK, NADH dehydrogenase I chain K, len: 101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase I chain K (EC 1.6.5.3) (102 aa), fasta scores; E(): 1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam match to entry PF00420 oxidored q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"
 /codon_start=1
 /transl_table=11
 /product="NADH dehydrogenase I chain K"
 /protein_id="CAB83325.1"
 /db_xref="GI:7378783"
 /db_xref="SPTREMBL:Q9JQU9"
 /translation="MITLTVLGLALLFGISAMGIFMNRKNVLVLLMSIELMLDAN
 ENFIASFQHLGDTAQGIFFVFVLTVAAESALGLAIMVLVYRNRTINNVADLDELKG"
 complement(3564..3863)
 /gene="nuoK"
 /note="Pfam match to entry PF00420 oxidored q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L, score 141.50, E-value 1.5e-38"

misc_feature

misc_feature

NADH-Ubiquinone/plastoquinone (complex I), various chains, score 361.80, E-value 7.3e-102"
 complement(2061..2243)
 /gene="nuoL"
 /note="Pfam match to entry PF00662 oxidored q1_N, NADH-Ubiquinone oxidoreductase (complex I), Chain 5 N-terminus, score 77.20, E-value 3.5e-19"
 complement(2487..2828)
 /gene="NMA0003"
 complement(2487..2828)
 /gene="NMA0003"
 /note="NMA0003, unknown, len: 113 aa"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein NMA0003"
 /protein_id="CAB83323.1"
 /db_xref="GI:7378781"
 /db_xref="SPTREMBL:Q9JX91"
 /translation="NMNKNLEQEPHKAMLIYQEALNLPQYKATRFLOIVNERGGK
 EADKXLSIGERKKTGTGTETLLSGGVHALKYSMEYLVQLKPMCDLFTFEQLAVARK
 RLKRVGFVSPK"
 complement(2833..2837)
 /gene="NMA0004"
 complement(2952..3527)
 /gene="NMA0004"
 complement(2952..3527)
 /gene="NMA0004"
 /note="NMA0004, unknown, len: 191 aa; similar to hypothetical proteins e.g. Y977_HAEIN HI977 (191 aa), fasta scores; E(): 0; 80.5% identity in 190 aa overlap, and to proteins involved in cell cycle e.g. TR:O54679 (EMBL:AF036487) Lactococcus lactis Plasmid pNZ4000 putative mobilization protein (200 aa), fasta scores; E(): 0.55.1% identity in 167 aa overlap, and FIC_ECOLI cell filamentation protein FIC (200 aa), fasta scores; E(): 0.093, 29.4% identity in 126 aa overlap"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein NMA0004"
 /protein_id="CAB83324.1"
 /db_xref="GI:7378782"
 /db_xref="SPTREMBL:Q9JQR9"
 /translation="MPSENPIKTKWKSIDESQLHNARLFPESGDIDRIEVTAGLQQ
 ITRYLFGGLYDPAGIREDNISKGGFRFANAYLKEALVKIEQMPERFETIIAKYVE
 MNIAHPFLGEGNRSTRIWLDVLKKNLKVNNQVNSKTYLQAMERSFVNDLELRFLL
 LKNDLITDDVNRREIIFKGIQESYVYEGYK"
 complement(3516..3525)
 /gene="NMA0004"
 /note="Core DNA uptake sequence: gccgtctgaa"
 /label=DUS
 complement(3561..3866)
 /gene="nuoK"
 complement(3561..3866)
 /gene="nuoK"
 /EC_number="1.6.5.3"
 /note="NMA0005, nuoK, NADH dehydrogenase I chain K, len: 101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase I chain K (EC 1.6.5.3) (102 aa), fasta scores; E(): 1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam match to entry PF00420 oxidored q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"
 /codon_start=1
 /transl_table=11
 /product="NADH dehydrogenase I chain K"
 /protein_id="CAB83325.1"
 /db_xref="GI:7378783"
 /db_xref="SPTREMBL:Q9JQU9"
 /translation="MITLTVLGLALLFGISAMGIFMNRKNVLVLLMSIELMLDAN
 ENFIASFQHLGDTAQGIFFVFVLTVAAESALGLAIMVLVYRNRTINNVADLDELKG"
 complement(3564..3863)
 /gene="nuoK"
 /note="Pfam match to entry PF00420 oxidored q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L, score 141.50, E-value 1.5e-38"

```

Gene      complement(3863..4534)
CDS       /gene="nuoJ"
          complement(3863..4534)
          /gene="nuoJ"
          /EC_number="1.6.5.3"
          /note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:
          223 aa; similar to many e.g. NUOJ, RHOCA NADH dehydrogenase
          I chain J (EC 1.6.5.3) (202 aa), Fatsa scores; E():
          9.7e-21; 41.3% identity in 201 aa overlap. Contains Pfam
          match to entry PF00499 oxidized q3.
          NADH-ubiquinone/plastoquinone oxidoreductase chain 6"
          /codon_start=1
          /transl_table=11
          /product="NADH dehydrogenase I chain J"
          /protein_id="CAB83326.1"
          /db_xref="GI:7378784"
          /db_xref="SPTREMBL:O9JX90"
          /translation="MTFQILFYIFAVILYGAIKTVYAKNPVAAHLHLVTFVCVSAM
          LMLMQAFLGVTVLVVGVAVNVLFFVVMMLNIDIEEMRAGFWRHAPVAGVVGTL
          AVALILVNPVKTDIAAFGLMKDIPADYNNIRDLGSRTYTDLLPFLFAAAILLGMV
          RAIALVHRKTVNPKMDPADQVKVRADQGRMLVKMEAVKQPTESAEESEVSDDLKPK
          BEGKA"
          complement(3874..3879)
          /gene="nuoJ"
          complement(4034..4525)
          /gene="nuoJ"
          /note="Pfam match to entry PF00499 oxidized q3,
          NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
          score 77.90, E-value 2e-19"
          complement(4570..5277)
          /gene="NMA0007"
          complement(4570..5277)
          /gene="NMA0007"
          /note="NMA0007, unknown, len: 235 aa"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein NMA0007"
          /protein_id="CAB83327.1"

Query Match      46.5%; Score 55.8; DB 1; Length 340806;
Best Local Similarity 79.8%; Pred. No. 1.5e-08;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 17 AGTATGAATGCTGAGGACGACTGTCTCGAAGCAAACTCTTACCG---AA 73
Db 17124 ATTAATGAATGCTGAGGACGACTGTCTCGAAGCAAACTCTTACCGCAA 17183

QY 74 GTCTTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 112
Db 17184 GACTTCTATACCAAGCTCAATAGCGGCTAAGGAGAGA 17222

RESULT 8
AX024133/c
LOCUS      AX024133      2094 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION      Sequence 76 from Patent FR2785293.
ACCESSION      AX024133
VERSION      AX024133.1
KEYWORDS
SOURCE      Neisseria gonorrhoeae
ORGANISM      Neisseria gonorrhoeae
              Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
              Neisseriaceae; Neisseria.
REFERENCE      1
AUTHORS      Nassif X., Tinsley C., Aujane L., Perrin A., Rokbi B.,
              Bouchardon A. and Renaud M.G.
JOURNAL      Patent: FR 2785293-A 76 05-MAY-2000;
              PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES
            Location/Qualifiers
            1..2094
              /organism="Neisseria gonorrhoeae"
              /mol_type="genomic DNA"
              /db_xref="taxon:485"

```

```

1..2094
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC08942.1"
/db_xref="GI:10184445"
/translation="MNSTASKTLKGLSVFFASGFCALIQVSWQRLLFSGHIGIDLS
ITVIISFVWGLGVGAYFGGRIADRPSSIIPLFCIAEVSIGLFGVSKGLISGLGHL
LVADLPITAAANFLLLPTFMWGATLLTCTFFNRKHNHVGESIGTLYFFNTLGA
LGSLAAEFYVFTLSOIALTACINLLIAASVCCVTERMDVNTKPTSVNMLSF
LTGLSLGIEVLWVNFSAQSVFQAFSLIACFLTGAGVAFGRICRSRPFVDIP
FIGCFLWAGIADFLGAAWLTGFGFVHHAGITILSAVVVRGLIFLPHVHVGTDG
NKGRQVSNVYFANVAGSALGPLVIGLIDLLSTQOYLILCLISAAVFLFCTLPQK
SLRINAVSVAVSLMFGILMFLPDSVFQNIAGRPDLRIENKHGIYAVYHGGKVVY
ANYVDGANTDIFNSVNGIERAYLLPSLKGIRIFEVGLSTGSWAVLSAIPMQSM
IVAEINPAYRSIIADPEIAPLLODKRVEIIVLDGKRLLRHPDERFDLILMNSTWY
RAYSTNLLGAEFLKOVQSHLTDPGIYMENTHSPAFATAVHSIPYAVYGHMVVGA
TPVFNKLLKQRLKRWPSGRHVFDSSIVDAAAKQVSRMLRMTPEPSAGAEVI
TDNNIVEYKIRGI"
BASE COUNT      428 a 495 c 515 g 656 t
ORIGIN
Query Match      45.8%; Score 55; DB 6; Length 2094;
Best Local Similarity 81.1%; Pred. No. 3.8e-08;
Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;

QY 21 AATGAAATCCCTGAGCAGGACTGTCTGCGAAGCAAACTCTTACCG---AAGTCT 77
Db 743 AATGAAATCCCTGAGCAGGACTGTCTGCGAAGCAAACTCTTACCGCAAGACT 684

QY 78 TCTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 112
Db 683 TCTATACCCAGGCTCAATAGCGGCTCAAGGAGAGA 649

RESULT 9
AE002381/c
LOCUS      AE002381      10467 bp      DNA      linear      BCT 25-MAY-2000
DEFINITION      Neisseria meningitidis serogroup B strain MC58 section 23 of 206 of
              the complete genome.
ACCESSION      AE002381
VERSION      AE002381.1
KEYWORDS      GI:7225455
SOURCE      Neisseria meningitidis MC58
ORGANISM      Neisseria meningitidis MC58
              Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
              Neisseriaceae; Neisseria.
REFERENCE      1 (bases 1 to 10467)
              Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C.,
              Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
              Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D.,
              Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
              Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E.,
              Citterone H., Clark B.B., Cotton M.D., Utterback T.R., Khouri H.,
              Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V.,
              Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R.,
              Rappuoli R. and Venter J.C.
              Complete genome sequence of Neisseria meningitidis serogroup B
              strain MC58
JOURNAL      Science 287 (5459), 1809-1815 (2000)
MEDLINE      20175755
PUBMED      10710307
REFERENCE      2 (bases 1 to 10467)
              Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C.,
              Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
              Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D.,
              Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
              Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E.,
              Citterone H., Clark B.B., Cotton M.D., Utterback T.R., Khouri H.,
              Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V.,
              Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R.,
              Rappuoli R. and Venter J.C.
              Direct Submission
TITLE

```

JOURNAL Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

FEATURES
source

1..10467
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
69..170
/gene="NMB0233"
69..170
/gene="NMB0233"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40688.1"
/db_xref="GI:7225456"
/translation="MGFFYESSPKQTAFIAIVDEAHATMYTQSK"
517..1056
/gene="NMB0234"
517..1056
/gene="NMB0234"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40689.1"
/db_xref="GI:7225457"
/translation="MSQVFKDFDLSSVMKTSWADENYKEAPFTPEILAAVESEIL
LPQSFILMAVQNGGIFVKNCPPTQRRSWAENHVQICVSGIGCFKEGSLCGAM
LWLEWNPYIGWYFANDPSGGHAMFALDIYRACGKGPKVVFVEQSDSEFVEILV
PETFIRSLRHEDEFIDEET"
complement(1297..1458)
/gene="NMB0235"
complement(1297..1458)
/gene="NMB0235"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40690.1"
/db_xref="GI:7225458"
/translation="MMIIINYKGGDIIILTREELQLLSLVIEIYACVCIADAEFTF
SGRNPKLYV"
complement(1466..1669)
/gene="NMB0236"
complement(1466..1669)
/gene="NMB0236"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40691.1"
/db_xref="GI:7225459"
/translation="MLKRLNGKDTFKLITFIPLRRLQSGHRAINWDDKSGAIVF
PNSKDGGAFTPLGTYFDKQK"
complement(1582..1941)
/gene="NMB0237"
complement(1582..1941)
/gene="NMB0237"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40692.1"
/db_xref="GI:7225460"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

```

"
5855. 6448
/gene="NMB0243"
5855. 6448
/gene="NMB0243"
/note=similar to GP.3282561 percent identity: 69.41;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase I, C subunit"
/protein_id="RAF40697.1"
/db_xref="GI:7225463"
/translation="MASIQDLYETVSRVLGNQAKVISAIGEIVTECLPHEYSVMTA
LRDEEHLFELLDGVYDYSTYKNEAQGRFAVQSLLSKVKNQIRVRYVWVSDDD
FPVSVSDYVNSADWVEREAFDYGIMFNHNPDLRLITDYGFGVHPFRKDFPISGY
VEMRYDEQKVIYQPVTEIPREITPRIVRENYGGQ"
6438. 7694
/gene="NMB0244"

Query Match 45.8%; Score 55; DB 1; Length 10467;
Best Local Similarity 81.1%; Pred. No. 3.5e-08;
Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;

QY 21 AATGAAATGCTGTAGGACGAGCTGTCTGCGAACGAAACATCTCTTACCG---AAGTCT 77
|||||
Db 3200 AATGAAATGCTGTAGGACGAGCTGTCTGCGAACGAAACATCTCTTACCGACCAAGACT 3141
|||||

QY 78 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 112
|||||
Db 3140 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 3106
|||||

RESULT 10
AX043922/c 349980 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION
Sequence 1 from Patent WO0066791.
ACCESSION
AX043922
VERSION
AX043922.1 GI:11342850
KEYWORDS
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1
Piazza M., Hickey E., Peterson J., Tettelin H., Venter J.C.,
Maignani V., Galeotti C., Mora M., Ratti G., Scarcelli M.,
Scarlato V., Rappuoli R., Frazer C.M. and Grandi G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 1 09-NOV-2000;
CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
Location/Qualifiers
1.349980
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces--seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 1200001 to 1549980
349980 bases--seq 112: 1500001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
2272325 172325 bases"
BASE COUNT 83241 a 85091 c 95206 g 86442 t
ORIGIN

Query Match 45.8%; Score 55; DB 6; Length 349980;
Best Local Similarity 81.1%; Pred. No. 2.8e-08;
Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;

QY 21 AATGAAATGCTGTAGGACGAGCTGTCTGCGAACGAAACATCTCTTACCG---AAGTCT 77
|||||
Db 244794 AATGAAATGCTGTAGGACGAGCTGTCTGCGAACGAAACATCTCTTACCGACCAAGACT 244735
|||||

"
78 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 112
|||||
Db 244734 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 244700
|||||

QY 78 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 112
|||||
Db 244734 TCTATACCCAGCTCAATAGCGCTCAAGGAGAGA 244700
|||||

RESULT 11
AC113844/c 229303 bp DNA linear HTG 19-NOV-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-134E24, *** SEQUENCING IN PROGRESS
*** 2 unordereged pieces.
ACCESSION
AC113844
VERSION
AC113844.6 GI:25072835
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 229303)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaru, A., Barker, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, W., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karkathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwono, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
Fuazo, M., Quitoz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
Sheed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Stemle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 229303)
Worley, K. C.
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE

```



```

Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meidriam,J., Meneus,L., Mihov,I., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo/A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 22, 2003 this sequence version replaced gi:28460968.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L28903
Center clone name: 266_K_11
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 186175 bases at least Q40
Consensus quality: 186660 bases at least Q30
Consensus quality: 186941 bases at least Q20
Insert size: 193000; agarose-fp
Insert coverage: 187586; sum-of-ctgts
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 12.4 in Q20 bases; sum-of-ctgts
-----
** NOTE: This is a 'working draft' sequence. It currently
** consists of 5 contigs. The true order of the pieces
** is not known and their order in this sequence record is
** arbitrary. Gaps between the contigs are represented as
** runs of N, but the exact sizes of the gaps are unknown.
** This record will be updated with the finished sequence
** as soon as it is available and the accession number will
** be preserved.
1
26328: contig of 26328 bp in length
26329 26428: gap of 100 bp
26429 28968: contig of 2540 bp in length
28969 29068: gap of 100 bp
29069 145708: contig of 116640 bp in length
145709 145808: gap of 100 bp
145809 173362: contig of 27554 bp in length
173363 173462: gap of 100 bp
173463 187986: contig of 14524 bp in length.
Location/Qualifiers
1..187986
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="13"
/map="13"
/clone="RP23-266K11"
/clone_lib="RPCI-23 Female Mouse BAC"
1..26328
/note="assembly_fragment
clone end:SP6
vector side:left"
ature

```

```

misc_feature      26429..28968
                    /note="assembly_fragment"
misc_feature      29069..145708
                    /note="assembly_fragment"
misc_feature      145809..173362
                    /note="assembly_fragment"
misc_feature      173463..187986
                    /note="assembly_fragment"
                    clone_end:77
                    vector_side:right"
BASE COUNT      55004 a 39677 c 38734 g 54171 t 400 others
ORIGIN
Query Match      25.2%; Score 30.2; DB 2; Length 187986;
Best Local Similarity 60.2%; Pred. No. 15;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1 CGGTGAGAAACAGGCAAGGTAATGAAATGCTGAGGACGAGTGTCTGCTCGAACGAAA 60
Db 61038 CAGTCAGGAAAGCAAAATGATGACATGAAACAGCAGCGAATCTCTCGAAGGAAA 61097
QY 61 ACTCTTACCGAAGTCTTCTATA 83
Db 61098 CATGCAACACATGTGATATACA 61120

RESULT 14
AC096226      254098 bp DNA linear HTG 10-MAY-2003
LOCUS      Rattus norvegicus chromosome 18 clone CH230-26M18, WORKING DRAFT
DEFINITION      SEQUENCE, 8 unordered pieces.
ACCESSION      AC096226
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS      Rattus norvegicus
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
REFERENCE      1 (bases 1 to 254098)
AUTHORS      Muzny, D., Marie, Metzkler, M., Lee, Abranzon, S., Adams, C., Alder, J.,
      Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
      Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
      Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
      Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
      Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
      Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
      Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
      Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
      Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
      Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
      Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
      Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
      Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
      Fraser, C. M., Gabisi, A., Gant, R., Garcia, A., Garner, T., Garza, M.,
      Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
      Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
      Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
      Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
      Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebirt, D., Jackson, A.,
      Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
      Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
      Kow, S. C., Kratt, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
      Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
      Lorensuhowa, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J.,
      Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
      Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
      Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
      Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
      Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
      Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
      Nwackemeloh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
      Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,

```

```

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Fuaro, M., Quiroz, J., Rachlin, E., Reeves, K., Resier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Speed, A., Sodargren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G., and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 254098)
Worley, K. C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 254098)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GERQ
Center clone name: CH230-26M18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 245379 bases at least Q40
Consensus quality: 246654 bases at least Q30
Consensus quality: 247940 bases at least Q20
Estimated insert size: 258267; sum-of-contigs estimation
Quality coverage: 15x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1 243470: contig of 243470 bp in length
* 243471 243570: gap of unknown length
* 243571 244710: contig of 1140 bp in length

```

* 244711 244810: gap of unknown length
* 244811 245858: contig of 1048 bp in length
* 245859 245958: gap of unknown length
* 245959 247049: contig of 1091 bp in length
* 247050 247149: gap of unknown length
* 247150 248930: contig of 1781 bp in length
* 248931 249030: gap of unknown length
* 249031 250613: contig of 1583 bp in length
* 250614 250713: gap of unknown length
* 250714 252352: contig of 1639 bp in length
* 252353 252452: gap of unknown length
* 252453 254098: contig of 1646 bp in length.

FEATURES

source
1. 254098
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="18"
/clone="CH230-26M18"
misc_feature
1. 1258
/note="wgs_end_extension
clone_end:Sp6"
3663. 4971
/note="wgs_end_extension
clone_end:Sp6"
6192. 7383
/note="wgs_end_extension
clone_end:Sp6"
8638. 9356
/note="clone boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH334077"
misc_feature
complement(243074..243440)
/note="clone boundary
clone_end:T7
site:EcoRI
end_sequence:BH334076"
BASE COUNT 63505 a 57265 c 56401 g 71601 t 5326 others
ORIGIN

Query Match 25.2%; Score 30.2; DB 2; Length 254098;
Best Local Similarity 62.7%; Pred. No. 15;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 46 GTGCTCGAGCAAACTCTTACCGAGTCTCTATACCCAGGCTCAATAGCCCTCAA 105
|||||
Db 118293 GTGCTTGGATCTAAACTCTGCTTATAGCATTATACCGAGTCTCTCTGAGCCA 118352
|||||
QY 106 GGAGAGAGCTATCAT 120
Db 118353 GAAGGAAGTCTCTT 118367
|||||

RESULT 15
AC137237/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
*** 2 unordered pieces.
ACCESSION AC137237
VERSION AC137237.1 GI:25073139
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 204173)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnham, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyte, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Naif, L.,
Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackeleme, O., Okwuonu, G., Olarinpunsagcon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rockett, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Woodson, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission

Unpublished
2 (bases 1 to 204173)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: KZKB
Center clone name: CH230-unknown
----- Summary Statistics

```
Assembly program: Phrap: version 0.990329
Consensus quality: 185526 bases at least Q40
Consensus quality: 189577 bases at least Q30
Consensus quality: 192542 bases at least Q20
Estimated insert size: 193275; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 202904: contig of 202904 bp in length
* 202905 203004: gap of unknown length
* 203005 204173: contig of 1169 bp in length.
FEATURES
    source
        1..204173
            location/Qualifiers
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-unknown"
            1..11072
            /note="wgs_contig"
    misc_feature
        BASE COUNT 53905 a 40373 c 40514 g 59594 t 9787 others
        ORIGIN
            Query Match      25.0%; Score 30; DB 2; Length 204173;
            Best Local Similarity 54.5%; Pred. No. 18;
            Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
            QY 9 AACAGGCAAGGTAATGAAATCCCTGAGCGACGGACTGTGTGGCAACGAAACTCCTTA 68
            Db 186037 AAAAAGCTATGTGTGAGGTGGTGATGCTTGGGCCAGTCTGTCAAGGAAGAGTTTACA 185978
            QY 69 CCGNAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATC 118
            Db 185977 CCAAGTCTTCTAATACGTGGATTGCTTTTCCCAATGGGAGAGTAATTTC 185928
```

Search completed: November 15, 2003, 03:33:48
Job time : 536.201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 77.4482 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-74

Perfect score: 120

Sequence: 1 CGGTGAGAACAGGCAAGGT.....CTCAAGGAGAGAGCTATCAT 120

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	120	19 AAV03582	Neisseria meningit
2	120	100.0	120	19 AAV03581	Neisseria meningit
3	120	100.0	120	21 AAA15387	Genomic fragment o
4	120	100.0	120	21 AAA15388	Genomic fragment o
C 5	55.8	46.5	1494	21 AAZ33687	Neisseria meningit
C 6	55.8	46.5	1524	25 ABZ39448	N. gonorrhoeae nuc
C 7	55.8	46.5	1524	25 ABZ41169	N. gonorrhoeae nuc
C 8	55.8	46.5	2094	21 AAA15321	DNA encoding a pol

C 9	55	45.8	1494	21 AAZ33685	Neisseria gonorrhe
C 10	55	45.8	1494	21 AAZ33686	Neisseria meningit
C 11	55	45.8	2094	21 AAA15333	DNA encoding a pol
C 12	55	45.8	72750	21 AAA81468	N. meningitidis pa
C 13	55	45.8	345980	21 AAF21544	Neisseria meningit
C 14	55	45.8	1437668	21 AAA81490	N. meningitidis B
C 15	28.2	23.5	341511	24 ABS55200	Genomic DNA encodi
C 16	28.2	23.5	506286	24 ABS55320	Human transporter
C 17	28	23.3	2598	24 ABZ13841	Arabidopsis thalia
C 18	28	23.3	2956	21 AAC50837	Arabidopsis thalia
C 19	28	23.3	2957	21 AAC37988	Arabidopsis thalia
C 20	27.2	22.7	464	22 AAS44639	Human full-length
C 21	27.2	22.7	2908	23 ABL29769	Drosophila melanog
C 22	27.2	22.7	5025	23 ABL29768	Drosophila melanog
C 23	27.2	22.7	24417	18 AAT97221	Pseudomonas aerugi
C 24	27.2	22.7	159400	24 ABQ88126	Human osteoblast d
C 25	26.8	22.3	1661	22 ABA20581	Human nervous syst
C 26	26.8	22.3	1661	22 ABA20582	Human nervous syst
C 27	26.8	22.3	2493	22 ABA20580	Human nervous syst
C 28	26.6	22.2	359	21 AAG67192	Human nervous syst
C 29	26.6	22.2	364	21 AAG67194	E. grandis alpha,a
C 30	26.6	22.2	394	25 ABX48531	E. grandis alpha,a
C 31	26.6	22.2	397	21 AAG67186	Bovine EST associa
C 32	26.6	22.2	457	21 AAG67191	E. grandis alpha,a
C 33	26.6	22.2	717	25 ABX08946	E. grandis alpha,a
C 34	26.6	22.2	6319	23 ABL18447	CDNA encoding huma
C 35	26.6	22.2	29555	23 ABL18446	Drosophila melanog
C 36	26.4	22.0	327	24 ABN76131	Drosophila melanog
C 37	26.4	22.0	781	21 AAF12631	Human ORF1078 cDNA
C 38	26.4	22.0	1221	14 AAQ46776	Aspergillus oryzae
C 39	26.4	22.0	1221	14 AAQ46776	Maize dwarf mosaic
C 40	26.4	22.0	2682	22 ABA20583	Human nervous syst
C 41	26.4	22.0	2687	22 AAK84955	Human immune/haema
C 42	26.4	22.0	2687	22 ABA20584	Human nervous syst
C 43	26.4	22.0	2687	22 AAK84957	Human immune/haema
C 44	26.4	22.0	2688	22 ABA20579	Human nervous syst
C 45	26.4	22.0	8543	18 AAT47073	Maize dwarf mosaic
					Pyrococcus abyssi

ALIGNMENTS

RESULT 1

AAV03582

ID AAV03582 standard; DNA; 120 BP.

XX AC AAV03582;

XX DT 22-OCT-1998 (first entry)

XX DE Neisseria meningitidis DNA sequence C24.

XX N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;

XX KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;

XX KW meningitis; ss.

XX OS Neisseria meningitidis.

XX PN WO9802547-A2.

XX PD 22-JAN-1998.

XX PF 11-JUL-1997; 97WO-FR01295.

XX PR 12-JUL-1996; 96FR-0008768.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PA (SMIK) SMITHKLINE BEECHAM.

XX PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;

XX PI Vinals C;

DR WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX Example 4; Page 117; 150pp; French.

XX AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, *frpA* or *C*,

CC *opc*, *porA*, *rotamase*, sequence IC1106, IGA protease, *pilC*, *pilC*,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC *lactamica*, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX Sequence 120 BP; 38 A; 30 C; 30 G; 22 T; 0 other;

SQ Query Match 100.0%; Score 120; DB 19; Length 120;

Best Local Similarity 100.0%; Pred. No. 1.2e-33;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAAACAGGCAAGTAAATGAAATGCTGAGGCACGAGCTGTCTCGAACGAAA 60

DB 1 CGGTGAGAAACAGGCAAGTAAATGAAATGCTGAGGCACGAGCTGTCTCGAACGAAA 60

QY 61 ACTCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

DB 61 ACTCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

RESULT 2

AAV03581

ID AAV03581 standard; DNA; 120 BP.

XX AC AAV03581;

XX 22-OCT-1998 (first entry)

XX *Neisseria meningitidis* DNA sequence C20.

XX *N. gonorrhoeae*; *N. lactamica*; chromosome Z2491; region 1; region 2;

XX region 3; pathogenicity; blood-brain barrier; diagnosis; infection;

XX meningitis; ss.

XX *Neisseria meningitidis*.

XX WO9802547-A2.

XX 22-JAN-1998.

XX 11-JUL-1997; 97WO-FR01295.

XX 12-JUL-1996; 96PR-0008768.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX (SMK) SMITHKLINE BEECHAM.

XX Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;

XX Vinals C;

XX WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX Example 4; Page 116; 150pp; French.

XX AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, *frpA* or *C*,

CC *opc*, *porA*, *rotamase*, sequence IC1106, IGA protease, *pilC*, *pilC*,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC *lactamica*, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX Sequence 120 BP; 38 A; 30 C; 30 G; 22 T; 0 other;

SQ Query Match 100.0%; Score 120; DB 19; Length 120;

Best Local Similarity 100.0%; Pred. No. 1.2e-33;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAAACAGGCAAGTAAATGAAATGCTGAGGCACGAGCTGTCTCGAACGAAA 60

DB 1 CGGTGAGAAACAGGCAAGTAAATGAAATGCTGAGGCACGAGCTGTCTCGAACGAAA 60

QY 61 ACTCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

DB 61 ACTCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

RESULT 3

AAV15387

ID AAV15387 standard; DNA; 120 BP.

XX AC AAV15387;

XX 04-SEP-2000 (first entry)

XX Genomic fragment of *Neisseria meningitidis* Z2491.

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

XX *Neisseria meningitidis*.

XX WO2000026375-A2.

XX 11-MAY-2000.

XX 28-OCT-1999; 99WO-FR02643.

XX 30-OCT-1998; 98PR-0013693.

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

XX Tinsley C, Perrin A;

XX WPI; 2000-365622/31.

XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic

XX or preventative vaccines and for diagnosis

XX Claim 1; Page 39; 187pp; French.

XX The present sequence represents a genomic fragment of *Neisseria*

XX meningitidis. The specification describes proteins that are specific

CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,
 CC or their antigenic fragments, are used in vaccines to treat or protect
 CC against *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequences are also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 CC note: the present sequence is included in Claim 1, but it is those
 CC sequences that do not include the present sequence that are actually
 CC claimed.

XX
 SQ Sequence 120 BP; 38 A; 30 C; 30 G; 22 T; 0 other;
 Query Match 100.0%; Score 120; DB 21; Length 120;
 Best Local Similarity 100.0%; Pred. No. 1.2e-33;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAAACAGGCAAGGTAAATGAAATGCTTACCGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 60
 Db 1 CGGTGAGAAACAGGCAAGGTAAATGAAATGCTTACCGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 60
 QY 61 ACTCCTTACCGAAGTCTTCTATACCGAGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 120
 Db 61 ACTCCTTACCGAAGTCTTCTATACCGAGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 120

RESULT 4
 AAA15388
 ID AAA15388 standard; DNA; 120 BP.
 AC AAA15388;
 XX
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Genomic fragment of *Neisseria meningitidis* 22491.
 XX
 KW Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 XX
 OS *Neisseria meningitidis*.
 PN WO200026375-A2.
 PD 11-MAY-2000.

PF 28-OCT-1999; 99WO-FR02643.
 XX
 XX 30-OCT-1998; 98FR-0013693.
 XX
 XX (INNR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 XX
 DR WPI; 2000-365622/31.
 XX

PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 PT or preventative vaccines and for diagnosis
 XX
 PS Claim 1; Page 39; 187pp; French.

XX The present sequence represents a genomic fragment of *Neisseria*
 CC meningitidis. The specification describes proteins that are specific
 CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,
 CC or their antigenic fragments, are used in vaccines to treat or protect
 CC against *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequences are also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 CC note: the present sequence is included in Claim 1, but it is those
 CC sequences that do not include the present sequence that are actually
 CC claimed.

XX Sequence 120 BP; 38 A; 30 C; 30 G; 22 T; 0 other;

Query Match 100.0%; Score 120; DB 21; Length 120;
 Best Local Similarity 100.0%; Pred. No. 1.2e-33;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAAACAGGCAAGGTAAATGAAATGCTTACCGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 60
 Db 1 CGGTGAGAAACAGGCAAGGTAAATGAAATGCTTACCGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 60
 QY 61 ACTCCTTACCGAAGTCTTCTATACCGAGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 120
 Db 61 ACTCCTTACCGAAGTCTTCTATACCGAGGCTCAATAGCCGCTCAAGGAGAGACTATCAT 120

RESULT 5
 AAZ53687/c
 ID AAZ53687 standard; DNA; 1494 BP.
 XX
 AC AAZ53687;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE *Neisseria meningitidis* ORF 402 partial DNA sequence SEQ ID NO:1323.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy; ds.
 XX
 OS *Neisseria meningitidis*.

PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR P-PSDE; AAY74925.

PT Novel *Neisseria* polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 7; Page 726; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX SQ Sequence 1494 BP; 321 A; 354 C; 373 G; 446 T; 0 other;
Query Match 46.5%; Score 55.8; DB 21; Length 1494;
Best Local Similarity 79.8%; Pred. No. 5.4e-10;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
OY 17 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCG---AA 73
Db 147 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCGCAA 88
OY 74 GTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGA 112
Db 87 GACTTCTATACCAAGCTCAATAGCCGCTAGGAAAGA 49
RESULT 6
ABZ39448
ID ABZ39448 standard; DNA; 1524 BP.
XX AC ABZ39448;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 3485.
XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB02069.
XX PR 12-FEB-2001; 2001GB-0003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizsa M, Masignani V, Monaci E;
XX DR WPI; 2003-058415/05.
XX DR P-PSDB; ABP78478.
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX PS medicament for treating or preventing N. gonorrhoeae infection -
XX Disclosure; Page 443-444; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX molecules of the invention.
XX SQ Sequence 1524 BP; 458 A; 386 C; 355 G; 325 T; 0 other;
Query Match 46.5%; Score 55.8; DB 25; Length 1524;
Best Local Similarity 79.8%; Pred. No. 5.4e-10;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
OY 17 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCG---AA 73
Db 1345 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCGCAA 1404
OY 74 GTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGA 112
Db 1405 GACTTCTATACCAAGCTCAATAGCCGCTAGGAAAGA 1443

RESULT 7
ABZ41169/c
ID ABZ41169 standard; DNA; 1524 BP.
XX AC ABZ41169;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 6927.
XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB02069.
XX PR 12-FEB-2001; 2001GB-0003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizsa M, Masignani V, Monaci E;
XX DR WPI; 2003-058415/05.
XX DR P-PSDB; ABP80199.
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX PS medicament for treating or preventing N. gonorrhoeae infection -
XX Disclosure; Page 690; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX molecules of the invention.
XX SQ Sequence 1524 BP; 325 A; 355 C; 386 G; 458 T; 0 other;
Query Match 46.5%; Score 55.8; DB 25; Length 1524;
Best Local Similarity 79.8%; Pred. No. 5.4e-10;
Matches 79; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
OY 17 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCG---AA 73
Db 180 AGTAAATGAAATGCTGAGCAGCGACTGTGCTGCGAACGAAATCCTTACCGCAA 121
OY 74 GTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGA 112
Db 120 GACTTCTATACCAAGCTCAATAGCCGCTAGGAAAGA 82
RESULT 8
AAA15321/c
ID AAA15321 standard; DNA; 2094 BP.
XX AC AAA15321;
XX DT 04-SEP-2000 (first entry)
XX DE DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX OS Neisseria meningitidis.
XX


```

PR 01-MAY-1998; 98US-0083758.
PR 21-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR P-PSDB; AAY74924.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics
XX
XX Claim 7; Page 724-725; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of
XX the invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the
XX manufacture of medicaments for treating or preventing infection due to
XX Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
XX presence of Neisseria bacteria, or to raise antibodies. They may also
XX be used to screen for agonists or antagonists, which may themselves
XX have use as antibacterial agents. The polynucleotides of the invention
XX may also be used in gene therapy protocols.
XX
XX Sequence 1494 BP; 319 A; 354 C; 373 G; 442 T; 6 other;
SQ
XX
XX Query Match 45.8%; Score 55; DB 21; Length 1494;
XX Best Local Similarity 81.1%; Pred. No. 1e-09;
XX Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;
XX
QY 21 AATGAAATGCTGAGGACGGACTGTGCTGCGAACGAAACTCTTACCG---AAGTCT 77
Db 143 AATGAAATGCTGAGGACGGACTGTGCTGCGAACGAAACTCTTACCGCAAGACT 84
XX
QY 78 TCTATACCCAGCTCAATAGCCGCTCAAGGAGAGA 112
Db 83 TCTATACCCAGCTCAATAGCCGCTTAGGAAAGA 49
XX
XX RESULT 11
XX AAAL5333/c
XX ID AAAL5333 standard; DNA; 2094 BP.
XX
XX AC AAAL5333;
XX
XX 04-SEP-2000 (first entry)
XX
XX DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
XX Neisseria gonorrhoeae.
XX
XX Key Location/Qualifiers
XX CDS 1..2094
XX /*tag= a
XX
XX WO2000026375-A2.
XX

```

```

PD 11-MAY-2000.
XX
XX 28-OCT-1999; 99WO-FR02643.
XX
XX 30-OCT-1998; 98FR-0013693.
XX
XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX (INMR ) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX
XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
XX Tinsley C, Perrin A;
XX
XX WPI; 2000-365622/31.
DR P-PSDB; AAY93303.
XX
XX New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX or preventative vaccines and for diagnosis
XX
XX Claim 4; Page 170-173; 187pp; French.
XX
XX The present sequence encodes a protein that is specific for pathogenic
XX strains of Neisseria. The polynucleotides, polypeptides, or their
XX antigenic fragments, are used in vaccines to treat or protect against
XX Neisseria infections, particularly by N. meningitidis. The
XX polynucleotide sequence is also used for recombinant production of
XX the polypeptide and to produce attenuated Neisseria strains that
XX overexpress it, or express it in a non-toxic mutant form.
XX
XX Sequence 2094 BP; 428 A; 495 C; 515 G; 556 T; 0 other;
SQ
XX
XX Query Match 45.8%; Score 55; DB 21; Length 2094;
XX Best Local Similarity 81.1%; Pred. No. 1.2e-09;
XX Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;
XX
QY 21 AATGAAATGCTGAGGACGGACTGTGCTGCGAACGAAACTCTTACCG---AAGTCT 77
Db 743 AATGAAATGCTGAGGACGGACTGTGCTGCGAACGAAACTCTTACCGCAAGACT 684
XX
QY 78 TCTATACCCAGCTCAATAGCCGCTCAAGGAGAGA 112
Db 683 TCTATACCCAGCTCAATAGCCGCTTAGGAAAGA 649
XX
XX RESULT 12
XX AAAL5333
XX ID AAAL5333 standard; DNA; 72750 BP.
XX
XX AC AAAL5333;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis partial DNA sequence gnm_16 SEQ ID NO:16.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX Meningococcus B; MenB; ds.
XX
XX Neisseria meningitidis.
XX
XX WO2000022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 09-OCT-1998; 98US-0103794.
XX 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
XX Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M,
XX Rappuoli R, Pizza M;
XX

```

```

XX WPI; 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea -
XX
XX Claim 7; Page 406-428; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. AA81453 to AA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences. AA82460 to AA8303 and AA82520 to AA82563 represent
XX Neisseria DNA sequences and their corresponding proteins; AA81254 to
XX AA81259 and AA81304 to AA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AA81322 to
XX AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
XX sequences, which are all used in the exemplification of the present
XX invention. The nucleic acid sequences, protein sequences, and antibodies
XX against them can be used in the manufacture of a composition. The
XX composition can be used as a medicament (or in the manufacture of a
XX medicament) for treating, preventing or diagnosing infection due to
XX Neisserial bacteria. For example, some of the identified proteins could
XX be components of vaccines against Meningococcus B; against all serotypes;
XX and/or against all pathogenic Neisseriae. Identification of sequences
XX from the bacterium will also facilitate production of biological probes,
XX particularly organism-specific probes. Attempts to make efficacious
XX Meningococcus B vaccines have failed mainly due to antigen tolerance.
XX Multivalent vaccines have also been tried but none have successfully
XX overcome antigenic variability. The provision of further, complete
XX sequences may provide an opportunity to identify secreted or surface
XX exposed proteins that may be presumed targets for the immune system and
XX which are not antigenically variable or at least more conserved than
XX other more variable regions.
XX
XX Sequence 72750 BP; 17518 A; 19945 C; 18810 G; 16477 T; 0 other;
XX
XX Query Match 45.8%; Score 55; DB 21; Length 72750;
XX Best Local Similarity 81.1%; Pred. No. 4e-09;
XX Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;
XX
XX QY 21 AATGAAATGCTGAGGCACGGACTGTGCGAAGCAAACTCCTTACCG---AAGTCT 77
XX DB 56066 AATGAAATGCTGAGGCACGGACTGTGCGAAGCAAACTCCTCACCACAGACT 56125
XX
XX QY 78 TCTATACCGAGCTCAATAGCCGCTCAAGGAGAGA 112
XX DB 56126 TCTATACCGAGCTCAATAGCCGCTCAAGGAGAGA 56160
XX
XX RESULT 13
XX AAF21544/C
XX ID AAF21544 standard; DNA; 349980 BP.
XX AC AAF21544;
XX
XX DT 13-MAR-2001 (first entry)
XX
XX DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.
XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
XX
XX OS Neisseria meningitidis.
XX
XX PN WO200066791-A1.
XX
XX XX 09-NOV-2000.
XX
XX PF 08-MAR-2000; 2000WO-US05928.
XX
XX XX 30-APR-1999; 99US-0132069.
XX
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 28-FEB-2000; 2000GB-0004895.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Piza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masiagnani V;
XX Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V, Rappuoli R;
XX Frazer CM, Grandi G;
XX
XX WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -
XX
XX Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
XX Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
XX to AAF21613 represent fragments of the NMB genomic sequence, as the
XX sequence was too long to go in a record on its own it was split into 8
XX sequences which overlap each other at the beginning and end of each
XX sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
XX the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
XX the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
XX Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
XX AAF21606 represent PCR primers which are used in the exemplification of
XX the present invention. The NMB genome and fragments from it have
XX antibacterial activity, and can be used in vaccines and gene therapy.
XX Neisseria nucleic acids, proteins and/or antibodies which binds to the
XX proteins can be used in compositions for treating or preventing infection
XX due to Neisserial bacteria or as a diagnostic reagent for detecting the
XX presence of Neisserial bacteria or of antibodies raised to Neisserial
XX bacteria. Computers, computer memory, computer storage medium or computer
XX databases can be used in a search to identify open reading frames (ORFs)
XX or coding sequences within the NMB genome. The DNA sequences provide
XX further opportunities to find antigenic or immunogenic proteins which are
XX more effective in vaccines than the outer membrane proteins currently
XX used.
XX
XX Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other;
XX
XX Query Match 45.8%; Score 55; DB 21; Length 349980;
XX Best Local Similarity 81.1%; Pred. No. 7e-09;
XX Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;
XX
XX QY 21 AATGAAATGCTGAGGCACGGACTGTGCGAAGCAAACTCCTTACCG---AAGTCT 77
XX DB 244794 AATGAAATGCTGAGGCACGGACTGTGCGAAGCAAACTCCTCACCACAGACT 244735
XX
XX QY 78 TCTATACCGAGCTCAATAGCCGCTCAAGGAGAGA 112
XX DB 244734 TCTATACCGAGCTCAATAGCCGCTCAAGGAGAGA 244700
XX
XX RESULT 14
XX AAF21490/C
XX ID AAF21490 standard; DNA; 1437668 BP.
XX AC AAF21490;
XX
XX DT 04-DEC-2000 (first entry)
XX
XX DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX Meningococcus B; MenB; ds.
XX
XX OS Neisseria meningitidis.
XX
XX PN WO200022430-A2.
XX

```

PD 20-APR-2000.
 XX
 XX 08-OCT-1999; 99WO-US23573.
 XX
 PR 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 XX (CHIR) CHIRON CORP.
 PA
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX
 DR WPI; 2000-318079/27.
 XX
 XX Isolated nucleotide sequences of *Neisseria meningitidis* which can be
 PT used in the diagnosis and treatment of *N. meningitidis* infection and
 PT other *Neisseria* infections, for example, *N. gonorrhoea*.
 XX
 PS Claim 7; Page 866-1272; 1760pp; English.
 XX
 CC The present invention describes methods of obtaining immunogenic
 CC proteins from *Neisseria* genomic sequences. AA81453 to AA82414
 CC represent specifically claimed *Neisseria meningitidis* genomic DNA
 CC sequences; AA81260 to AA81303 and AA825620 to AA825663 represent
 CC *Neisseria* DNA sequences and their corresponding proteins; AA81254 to
 CC AA81259 and AA81304 to AA81321 represent PCR primers used in the
 CC isolation of *Neisseria meningitidis* DNA sequences; and AA81322 to
 CC AA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to a
 CC *Neisseria* bacteria. For example, some of the identified proteins could
 CC be components of vaccines against *Meningococcus B*; against all serotypes;
 CC and/or against all pathogenic *Neisseria*. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC *Meningococcus B* vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 XX
 SQ Sequence 1437668 BP; 344338 A; 353206 C; 385074 G; 355045 T; 5 other;
 Query Match 45.8%; Score 55; DB 21; Length 1437668;
 Best Local Similarity 81.1%; Pred. No. 1.1e-08;
 Matches 77; Conservative 0; Mismatches 15; Indels 3; Gaps 1;
 QY 21 AATGAAATGCTGAGCAGCGACTGTGCTGCGAAGCAAACTCCTTACCG---AAGTCT 77
 Db 244794 AATGAAATGCTGAGCAGCGACTGTGCTGCGAAGCAAACTCCTTACCGCAACGACT 244735
 QY 78 TCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGA 112
 Db 244734 TCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGA 244700
 RESULT 15
 ABS55200
 ID ABS55200 standard; DNA; 341511 BP.
 XX
 AC ABS55200;
 XX
 DT 06-JAN-2003 (first entry)
 XX
 DE Genomic DNA encoding human transporter protein.
 XX Human; transporter protein; hypoglycaemia; antidiabetic; gene therapy;
 KW gene; ds; single nucleotide polymorphism; SNP.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 2001..339512
 FT /tag= a
 FT /products "Human transporter protein"
 FT 2001..2073
 FT /tag= b
 FT /number= 1
 FT 2074..30237
 FT /tag= c
 FT /number= 1
 FT 30238..30361
 FT /tag= d
 FT /number= 2
 FT 30362..40719
 FT /tag= e
 FT /number= 2
 FT 40720..40826
 FT /tag= f
 FT /number= 3
 FT 40827..45732
 FT /tag= g
 FT /number= 3
 FT 45733..45861
 FT /tag= h
 FT /number= 4
 FT 45862..59818
 FT /tag= i
 FT /number= 4
 FT 59819..59934
 FT /tag= j
 FT /number= 5
 FT 59935..65742
 FT /tag= k
 FT /number= 5
 FT 65743..66135
 FT /tag= l
 FT /number= 6
 FT 66136..96458
 FT /tag= m
 FT /number= 6
 FT 96459..96885
 FT /tag= n
 FT /number= 7
 FT 96886..197167
 FT /tag= o
 FT /number= 7
 FT 197168..197367
 FT /tag= p
 FT /number= 8
 FT 197368..244442
 FT /tag= q
 FT /number= 8
 FT 244443..244695
 FT /tag= r
 FT /number= 9
 FT 244696..267106
 FT /tag= s
 FT /number= 9
 FT 267107..267303
 FT /tag= t
 FT /number= 10
 FT 267304..338564
 FT /tag= u
 FT /number= 10
 FT 338565..339509
 FT /tag= v
 FT /number= 11
 FT replace (2557,T)
 FT /tag= w
 FT variation

```
FT variation /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (5492,C)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (5330,A)
FT /*tag= y
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (7269,C)
FT /*tag= z
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (9650,G/T)
FT /*tag= aa
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (10664,T)
FT /*tag= ab
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (11043,A/G/C)
FT /*tag= ac
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (13109,T)
FT /*tag= ad
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (13392,A)
FT /*tag= ae
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (13395..13397,GA)
FT /*tag= af
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (13563,C)
FT /*tag= ag
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (13694,A)
FT /*tag= ah
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (14075,A)
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (14573,A)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (15013,T)
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (15073..15075,GT)
FT /*tag= al
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (15848,T)
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (16346,T)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (16367,A)
FT /*tag= ao
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (16639,G)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (18149,A/T)
FT /*tag= aq
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (18243,C/A)
FT /*tag= ar
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (18293,G/A)
FT /*tag= as
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (19185,C)
FT /*tag= at
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (24728..24730,AA)
FT /*tag= au
FT /standard_name= "Single nucleotide polymorphism (SNP)"

FT variation replace (24790,T)
FT /*tag= av
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (24950,T)
FT /*tag= aw
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (28608,G)
FT /*tag= ax
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (29244,C)
FT /*tag= ay
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (29651,T)
FT /*tag= az
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (34034,T)
FT /*tag= ba
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (36671,G)
FT /*tag= bb
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (36836,T)
FT /*tag= bc
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (37010,A)
FT /*tag= bd
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (38461,A/C)
FT /*tag= be
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (38588,A/C)
FT /*tag= bf
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (39673,A)
FT /*tag= bg
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (39854,C)
FT /*tag= bh
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (41325,A)
FT /*tag= bi
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace (41412,A)
FT /*tag= bj
FT /standard_name= "Single nucleotide polymorphism (SNP)"

Query Match 23.5%; Score 28.2; DB 24; Length 341511;
Best Local Similarity 55.7%; Pred. No. 36;
Matches 54; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 21 AATCAAAATCCTGAGGCACGACTGTGCTGCGAAGCAAACTCCTTACCGAGTCTTCT 80
DB 193970 AGTCATTGGCTGTTGGAGAGAGGGTCTATAGACTGAGTTCTTTTCGGATGCTTAG 194029

QY 81 ATACCCAGGTCATATACCGCTCAAGGAGAGAGCTAT 117
DB 194030 ATCCAAATAATGATTAGACATCCCAAGAGATAGGTAT 194066

Search completed: November 15, 2003, 00:35:24
JOB time : 81.4482 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 800.53 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-74

Perfect score: 120

Sequence: 1 CGGTCAAGACAGGCAAGCT.....CTCAAGGAGAGACTATCAT 120

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31.8	26.5	545	9 AW669802	AW669802 113373 MA
C 2	31.2	26.0	589	12 BP009535	BP009535 BP009535
C 3	31.2	26.0	657	12 BP009682	BP009682 BP009682
4	30.4	25.3	479	29 FR0024955	AL017805 F.rubripe

C 5	30.2	25.2	711	10 BF859869	BF859869 963014H01
C 6	30.2	25.2	736	10 BF859841	BF859841 963014G04
C 7	30.2	25.2	782	10 BF859840	BF859840 963014G04
C 8	30.2	25.2	797	10 BF859868	BF859868 963014H01
C 9	30	25.0	1035	29 BZ558760	BZ558760 PA98401_4
C 10	29.4	24.5	966	10 BF127560	BF127560 601810209
C 11	29.2	24.3	771	29 BZ966854	BZ966854 PUGK45TB
C 12	29.2	24.3	816	13 BU745275	BU745275 CHI#008_E
C 13	29	24.2	768	12 BI218807	BI218807 602938279
C 14	28.8	24.0	571	29 FR0026279	FR0026279 F.rubripe
C 15	28.6	23.8	725	28 AQ960052	AQ960052 LERFB11TR
C 16	28.6	23.8	733	28 AQ960051	AQ960051 LERFB11TR
C 17	28.4	23.7	326	29 FR0024950	FR0024950 F.rubripe
C 18	28.4	23.7	538	14 CB886149	CB886149 3529_1_92
C 19	28.4	23.7	601	28 BH814982	BH814982 AM_Ba001
C 20	28.4	23.7	665	29 AG127283	AG127283 Pan trogl
C 21	28.4	23.7	691	12 BJ321620	BJ321620 BJ321620
C 22	28.4	23.7	693	12 BM865214	BM865214 mgapo10XP
C 23	28.4	23.7	723	12 BJ320632	BJ320632 BJ320632
C 24	28.2	23.5	597	9 AV593758	AV593758 AV593758
C 25	28	23.3	460	9 AA981455	AA981455 ua35f08.r
C 26	28	23.3	479	28 AZ048801	AZ048801 GSBBr0019
C 27	28	23.3	498	29 BZ603792	BZ603792 WHAAQ68TF
C 28	28	23.3	559	29 BZ606580	BZ606580 WHABB42TF
C 29	28	23.3	570	9 AL695123	AL695123 AL695123
C 30	28	23.3	676	28 BZ052416	BZ052416 jnr68f01
C 31	28	23.3	917	29 CC414833	CC414833 PUHPP45TB
C 32	28	23.3	980	29 CC213462	CC213462 CH261_115
C 33	28	23.3	1074	13 BU333735	BU333735 603869750
C 34	28	23.3	1278	13 BQ436845	BQ436845 AGENCOURT
C 35	27.8	23.2	234	29 FR0026265	FR0026265 F.rubripe
C 36	27.8	23.2	360	13 BY071431	BY071431 BY071431
C 37	27.8	23.2	400	10 BG101744	BG101744 EM1_5_G05
C 38	27.8	23.2	404	29 FR0023834	FR0023834 F.rubripe
C 39	27.8	23.2	535	10 BG464307	BG464307 EM1_72_B0
C 40	27.8	23.2	553	9 AI866077	AI866077 wl25d07.x
C 41	27.8	23.2	594	29 FR0031004	FR0031004 F.rubripe
C 42	27.8	23.2	707	12 BM683197	BM683197 UI-E-E01-
C 43	27.8	23.2	1181	13 BU253236	BU253236 603415556
C 44	27.6	23.0	472	29 CN300070	CN300070 Arabidops
C 45	27.6	23.0	482	14 H67236	H67236 yu65f05.r1

ALIGNMENTS

```

RESULT 1
1. AW669802 545 bp mRNA linear EST 25-APR-2001
LOCUS 113373 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AW669802
ACCESSION AW669802
VERSION AW669802.1 GI:7526316
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 545)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Partea,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366

```

TITLE

JOURNAL MEDLINE
PUBMED
COMMENT

Fax: 402 762 4390
 Email: smithemall.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACAT
 BACKWARD: GTTTCCTCAGTCAGCAG
 Plate: 108 row: E column: 18
 Seq primer: ATTAGTGACACTATAG.
 Location/Qualifiers

FEATURES

1..545
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 1BOV"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."
 130 a 142 c 116 g 157 t

BASE COUNT

ORIGIN

Query Match 26.5%; Score 31.8; DB 9; Length 545;
 Best Local Similarity 54.8%; Pred. No. 6.8;
 Matches 63; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 5 CAGAACAGCAGTAAATGCTGAGCAGCAGTGTGCTGCGAAGAACTC 64
 |||||
 DB 125 CAGTCACAGGGAGGAGATGCAATCTTGAGCAGTGTGGTGTGCTACCAACA 184

QY 65 CTACCGAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGCTATCA 119
 |||||
 DB 185 ATCAACGGTCTCATTTTTCAGTCTCTAAAGTGTGCAACAAAGAGTGTCTCA 239

RESULT 2

BP009535/c 589 bp mRNA linear EST 15-MAR-2002
 LOCUS BP009535 Nori Satoh unpublished cDNA library, young adult Ciona
 DEFINITION intestinalis cDNA clone c1ad71d03 5', mRNA sequence.

ACCESSION BP009535.1 GI:19501012
 VERSION
 KEYWORDS EST.

SOURCE

ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

REFERENCE 1 (bases 1 to 589)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

1..589
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="c1ad71d03"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young
 adult"
 171 a 130 c 111 g 177 t

BASE COUNT

ORIGIN

Query Match 26.0%; Score 31.2; DB 12; Length 589;
 Best Local Similarity 66.2%; Pred. No. 11;
 Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 27 AATGCTGAGCAGCAGTGTGCTGCGAAGAACTCTTACCGAGTCTTCTATACCC 86
 |||||
 DB 94 AAAGCTTGAGGAACGAAGTGTGTGTGAACCTAAATGCTTACCTATGGCAGCAACATCG 35

QY 87 AGGCTCAA 94
 |||||
 DB 34 AGCCTCAA 27

RESULT 3

BP009682/c

LOCUS BP009682 Nori Satoh unpublished cDNA library, young adult Ciona

DEFINITION intestinalis cDNA clone c1ad71n03 5', mRNA sequence.

ACCESSION BP009682

VERSION BP009682.1 GI:19501159

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

REFERENCE 1 (bases 1 to 657)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: sato@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1..657

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="c1ad71n03"

/tissue_type="whole animal"

/dev_stage="young adult"

/clone_lib="Nori Satoh unpublished cDNA library, young
 adult"

189 a 144 c 121 g 203 t

BASE COUNT

ORIGIN

Query Match 26.0%; Score 31.2; DB 12; Length 657;
 Best Local Similarity 66.2%; Pred. No. 12;
 Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 27 AATGCTGAGCAGCAGTGTGCTGCGAAGAACTCTTACCGAGTCTTCTATACCC 86
 |||||
 DB 124 AAAGCTTGAGGAACGAAGTGTGTGTGAACCTAAATGCTTACCTATGGCAGCAACATCG 65

QY 87 AGGCTCAA 94
 |||||
 DB 64 AGCCTCAA 57

RESULT 4

FR0024965

LOCUS FR0024965 479 bp DNA linear GSS 10-DEC-1997

DEFINITION F.rubripes GSS sequence, clone 135A08a66, genomic survey sequence.

ACCESSION AL017805

VERSION AL017805.1 GI:2684173

KEYWORDS GSS; genome survey sequence.

SOURCE Takifugu rubripes

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
1. (bases 1 to 479)
Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrana, Y.,
Williams, G. and Brenner, S.
Direct Submission
Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
Vector: pBluescript II KS
V-type: phagemid
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.

FEATURES

Location/Qualifiers

1..479
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="135A08a6"
/clone_lib="cosmid 135A08"
84 a 136 c 141 g 78 t 38 others

BASE COUNT

ORIGIN

Query Match 25.3%; Score 30.4; DB 29; Length 479;
Best Local Similarity 75.5%; Pred. No. 19;
Matches 37; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 61 ACTCCTACGAGTCTTCTATACCCAGGCTCAATAGCGGCTCAGGAG 109
Db 165 AGTCCATGAGAGCGCTTTTATCCAGGNTCACTGCCTCTGAAAGAG 213

RESULT 5

BF859869/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Charles Hauser

DCMB Box 91000

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1..711

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, Stress condition 1,

normalized, Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:

XhoI. This library, constructed by John Davies and Jeffrey

McDermott, combines cDNAs from CC-1690 cells grown to

mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,

1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr

, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was

purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoRI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT

ORIGIN

Query Match 25.2%; Score 30.2; DB 10; Length 711;
Best Local Similarity 60.2%; Pred. No. 26;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 7 GAAACAGCAAGGTAATGAAATCGTCAGCAGCGACTGTGTCGGAACGAAACCTCT 66
Db 178 GAAACTAGGAGCAATGACAAAGGAAAGTGAGGAGCGACGTTGGGAAACAAAGCCTCG 119
QY 67 TACCGAGTCTTCTATACCCAGG 89
Db 118 TCCGAAATTCCTGAGCCCGGGG 96

RESULT 6

BF859841/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Charles Hauser

DCMB Box 91000

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1..736

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, Stress condition I,

normalized, Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:

XhoI. This library, constructed by John Davies and Jeffrey

McDermott, combines cDNAs from CC-1690 cells grown to

mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,

1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr

, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was

purified from each sample, pooled and cDNA synthesized.

The cDNA was directionally cloned into lambda Zap II

(Stratagene) in the EcoRI (5') and XhoRI (3') sites.

pBluescript II SK- plasmids were excised from the lambda

ZAP clones by superinfection with ExAssist (Stratagene)

phage. The library was normalized using method 4 described

in Bonaldo et al (1996) Genome Research 6: 791-806."

176 a 195 c 216 g 149 t

BASE COUNT

ORIGIN

Query Match 25.2%; Score 30.2; DB 10; Length 736;
Best Local Similarity 60.2%; Pred. No. 27;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 7 GAACAGGCAAGTATGAATAATGCTGAGGACGAGTGTCTGCGAACCAAACTCT 66
|||||
DB 180 GAACTAGCGAAGTATGACAGGCGAAAGTGGGAGCAAAAGCTCTG 121
|||||

QY 67 TACCGAAGTCTTCTATACCCAGG 89
|||||
DB 120 TCCCGAATCTCTGACCCCGGG 98
|||||

RESULT 7
BF859840/c
LOCUS
DEFINITION
963014GQ4.v1 C. reinhardtii CC-1690, Stress condition I, normalized
, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BF859840
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
1 (bases 1 to 782)
Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C.,
Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 3
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

JOURNAL
COMMENT
Location/Qualifiers
1..782
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into Lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the Lambda
Zap clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 182 a 213 c 229 g 157 t
ORIGIN

Query Match 25.2%; Score 30.2; DB 10; Length 782;
Best Local Similarity 60.2%; Pred. No. 28;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 7 GAACAGGCAAGTATGAATAATGCTGAGGACGAGTGTCTGCGAACCAAACTCT 66
|||||
DB 175 GAACTAGCGAAGTATGACAGGCGAAAGTGGGAGCAAAAGCTCTG 116
|||||

QY 67 TACCGAAGTCTTCTATACCCAGG 89
|||||

Db 115 TGCGGAATCTCTGACCCCGGG 93

RESULT 8
BF859868/c
LOCUS
DEFINITION
963014H01.v1 C. reinhardtii CC-1690, Stress condition I, normalized
, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BF859868
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
1 (bases 1 to 797)
Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C.,
Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 3
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

JOURNAL
COMMENT
Location/Qualifiers
1..797
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into Lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the Lambda
Zap clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 188 a 215 c 229 g 160 t
ORIGIN

Query Match 25.2%; Score 30.2; DB 10; Length 797;
Best Local Similarity 60.2%; Pred. No. 28;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 7 GAACAGGCAAGTATGAATAATGCTGAGGACGAGTGTCTGCGAACCAAACTCT 66
|||||
DB 179 GAACTAGCGAAGTATGACAGGCGAAAGTGGGAGCAAAAGCTCTG 120
|||||

QY 67 TACCGAAGTCTTCTATACCCAGG 89
|||||
DB 119 TGCGGAATCTCTGACCCCGGG 97
|||||

RESULT 9
BZ558760/c
LOCUS
DEFINITION
BZ558760 1035 bp DNA linear GSS 17-DEC-2002
p984401_461.s1 pacs2-164 Pseudomonas aeruginosa genomic clone
p984401_461, genomic survey sequence.
BZ558760
ACCESSION
VERSION
BZ558760.1 GI:27173873

```

KEYWORDS      GSS.
SOURCE         Pseudomonas aeruginosa
ORGANISM       Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
REFERENCE      1 (bases 1 to 1035)
AUTHORS        Spencer D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
                Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE          Whole-Genome-Sequence variation among multiple isolates of
                Pseudomonas aeruginosa library
JOURNAL        J. Bacteriol., (2002) In press
COMMENT        Contact: Chris K. Raymond
                Genome Center
                University of Washington
                Box 352145, Seattle, WA 98105-2145, USA
                Tel: 2062216954
                Fax: 2066857244
                Email: craymond@u.washington.edu
                Class: shotgun.

FEATURES       Location/Qualifiers
                1..1035
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="2-164"
                /db_xref="taxon:287"
                /clone="pa98401.461"
                /clone_lib="pacs2-164"
                /notes="clinical isolate 2-164 whole genomic shotgun
                library."
BASE COUNT    168 a 337 c 333 g 196 t 1 others
ORIGIN
Query Match      25.0%; Score 30; DB 29; Length 1035;
Best Local Similarity 57.4%; Pred. No. 36;
Matches 54; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 19 GTAATGAATGCTCAGACCGACTGTGCTCGAACGAAACTCTTACCGAAGTCTT 78
Db 920 GAAGTGTGAATGTTGATACAGGCCGCTGGTGAATGCTTCGCTCAAGGAAGTGT 861

QY 79 CTATACCCAGGGTCAATAGCCGCTCAAGGAGAGA 112
Db 860 CGACGCCCGCCCTTGGCAGCGCTGCGCGGA 827

RESULT 10
BF127560
LOCUS
DEFINITION      601810209F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4053027 5',
                mRNA sequence.
ACCESSION      BF127560
VERSION        BF127560.1 GI:10966600
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished
                Contact: Robert Strausberg, Ph.D.
                Email: cga@bs-r@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA library Preparation: Ling Hong/Rubin Laboratory
                DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
                Clone Distribution by: Incyte Genomics, Inc.
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLC891 row: b column: 04
                High quality sequence stop: 393.
                Location/Qualifiers

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1..966
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4053027"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT    313 a 245 c 260 g 148 t
ORIGIN
Query Match      24.5%; Score 29.4; DB 10; Length 966;
Best Local Similarity 60.8%; Pred. No. 56;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 9 AACAGCAAGGTAATGAATGCTCGAGCGACTGTGCTCGAACGAAACTCTCTTA 68
Db 653 AACAGCGGGCGACAGAAACCCGCGGAACACACAGACGCGGGAGACGAAACGCCA 712

QY 69 CCGAAGTCTTCTATACCCA 87
Db 713 CAGAAGTCGACCAGACCAA 731

RESULT 11
BZ966854/c
LOCUS
DEFINITION      BZ966854 Zea mays genomic clone ZMMB7a378H17,
                genomic survey sequence.
ACCESSION      BZ966854
VERSION        BZ966854.1 GI:29184034
KEYWORDS       GSS.
SOURCE         Zea mays
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 771)
AUTHORS        Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
                A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
TITLE          Maize Genomics Consortium
JOURNAL         Unpublished
COMMENT         Other GSSs: PUGGK45TD
                Contact: Cathy Whitelaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whitelaw@tigr.org
                Seq primer: TR
                Class: sheared ends.
                Location/Qualifiers
                1..771
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMB7a378H17"
                /clone_lib="ZM_0.6_1.0_KB"
                /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"
BASE COUNT    204 a 179 c 169 g 219 t
ORIGIN

```

Query Match 24.3%; Score 29.2; DB 29; Length 771;
 Best Local Similarity 57.8%; Pred. No. 59;
 Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 6 AGAAACAGCAGGTAATGAAATGCTGAGCAGGACTGTGCTGCGAAGCAAACTCC 65
 |||||
 DB 752 AAAAAGTGGCTGAGAAACAAACACCATGGGCAACATGAGATGAGCAGTAAATCCA 693
 |||||

QY 66 TTACCGAAGTCTTCTATACCCAGGCTCAAT 95
 |||||
 DB 692 GTACCCAAAGTTTCAATACTAGAAGCAAT 663
 |||||

RESULT 12
 BU745275/c

LOCUS BU745275 816 bp mRNA linear EST 10-OCT-2002
 DEFINITION CH1#008_E047 Canine heart non-normalized cDNA Library in pBluescript Canis familiaris cDNA clone CH1#008_E04 5', mRNA sequence.

ACCESSION BU745275.1 GI:23694511
 VERSION BU745275
 KEYWORDS
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris

REFERENCE 1 (bases 1 to 816)
 AUTHORS Vi.Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.
 TITLE Expressed sequence tags from Canine heart
 JOURNAL Unpublished
 COMMENT Other ESTs: CH1#008_E0473
 Contact: George AL
 Division of Genetic Medicine
 Vanderbilt University
 529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
 Tel: 615 936 2660
 Fax: 615 936 2661
 Email: al.george@vanderbilt.edu
 Insert Length: 1116 Std Error: 0.00
 Seq primer: T7: TAATACGACTCACTATAGG
 High quality sequence start: 93
 High quality sequence stop: 419.
 Location/Qualifiers
 1..816
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /clone="CH1#008_E04"
 /tissue_type="heart"
 /cell_type="heart"
 /dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
 /clone_lib="Canine heart non-normalized cDNA Library in pBluescript"
 /notes="Orgn: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal) right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

BASE COUNT 171 a 250 c 214 g 180 t
 ORIGIN

Query Match 24.3%; Score 29.2; DB 13; Length 816;
 Best Local Similarity 59.8%; Pred. No. 61;
 Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 25 AAAATGCTGAGCAGGAGTGTGCTGCGAAGCAAAACTCTTACCGAAGTCTTCTATAC 84
 |||||
 DB 265 AAAATCCCGAGCTCAGTGGCGGAGCGGAAACACACCTACCGAAGTCTTCTGAT 206
 |||||

QY 85 CCAGGCTCAATACCGCTCAAG 106
 |||||

Db 205 GGAGCCTCGGAGGCGGAAAAG 184

RESULT 13
 BI218807

LOCUS BI218807 768 bp mRNA linear EST 11-JUL-2001
 DEFINITION 602938279F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5101627 5', mRNA sequence.

ACCESSION BI218807.1 GI:14672251
 VERSION BI218807
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 768)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: L1AM11244 row: m column: 20
 High quality sequence stop: 766.
 Location/Qualifiers
 1..768
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5101627"
 /lab_host="NCI_CGAP_Li9"
 /clone_lib="NCI_CGAP_Li9"
 /notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 Technolgies. Note: this is a NCI_CGAP Library."

BASE COUNT 163 a 202 c 215 g 188 t
 ORIGIN

Query Match 24.2%; Score 29; DB 12; Length 768;
 Best Local Similarity 57.0%; Pred. No. 69;
 Matches 53; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 13 GGCAAGGTAATGAAATGCTGAGGACGAGCTGTGCTGCGAAGCAAACTCTTACCGA 72
 |||||
 DB 493 GGCTGGCGAAGAAATCACTGGAGGACTGGTGACGAGGCGGACGACACCTGTGA 552
 |||||

QY 73 AGTCTTATATACCCAGGCTCAATAGCGGCTCAA 105
 |||||
 DB 553 TGCCTTACCGCCGAGGCTGGGCGAGCCCATCAA 585
 |||||

RESULT 14
 FR0026279

LOCUS FR0026279 571 bp DNA linear GSS 10-DEC-1997
 DEFINITION F.rubripes GSS sequence, clone 169K01ac9, genomic survey sequence.
 ACCESSION AL019112
 VERSION AL019112.1 GI:2685480
 KEYWORDS GSS: genome survey sequence.
 SOURCE Takifugu rubripes
 ORGANISM Takifugu rubripes

REFERENCE 1 (bases 1 to 571)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.

AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@gmp.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V type: phagemid
PRIMER: KS
DESCR: One pass dye-terminator sequencing of cosmid cloned genomic
sequence.

FEATURES

Location/Qualifiers

1..571
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="169K01aC9"
/clone_lib="cosmid 169K01"

BASE COUNT 97 a 166 c 158 g 92 t 58 others

ORIGIN

Query Match 24.0%; Score 28.8; DB 29; Length 571;
Best Local Similarity 72.0%; Pred. NO. 71;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 59 AAATCCTTACCGAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGA 108
Db 39 ANAGTCCATGCAGAGCGCTTCTTACCCAGGCTCAGCTCTCTGAAGA 88

RESULT 15
AQ960052/c

LOCUS

DEFINITION LERPB11TR LERA Arabidopsis thaliana genomic clone LERPB11, genomic
survey sequence.

ACCESSION

VERSION AQ960052.1 GI:6787876

KEYWORDS

GSS.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,
Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.

TITLE

Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based

JOURNAL

COMMENT

polymorphisms
Unpublished
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at@tigr.org
For additional information, see <http://www.tigr.org/tdb/at/at.html>
Seq primer: TR
Class: shotgun.

FEATURES

source

Location/Qualifiers

1..725
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERPB11"
/clone_lib="LERA"
/notes="Organ: Leaf; Vector: pHS01; Total genomic DNA was
sheared to 0.9-1 Kbp before ligation."

BASE COUNT

236 a 133 c 169 g 187 t

ORIGIN

Query Match

23.8%; Score 28.6; DB 28; Length 725;

Best Local Similarity 61.3%; Pred. No. 92;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 13 GGCAAGGTAATGAAAATGCCTGAGGACGAGCTGTGCGAAACGAAAACCTCTTACCGA 72
Db 602 GGCAGGTTAGTGTAAATTTCTAACCTGTTGAGTATCTTGAGAACTAAAGCATCTCTCAA 543
QY 73 AGCTTCTATACCCA 87
Db 542 AATCTTCTGTAACA 528

Search completed: November 15, 2003, 08:02:24
Job time : 803.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 18.691 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-74

Perfect score: 120

Sequence: 1 CGGTAGACAGGCAAGGT.....CTCAAGGAGAGAGCTATCAT 120

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A-COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B-COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A-COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B-COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PTUS-COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27.2	22.7	24317	2	US-08-846-762-1
2	26.4	22.0	1821	1	US-08-229-287-3
3	26.4	22.0	1821	1	US-08-496-944-1
4	26.2	21.8	5543	2	US-08-687-080-101
5	25.4	21.2	657	3	US-09-385-982-326
6	25.4	21.2	5303	4	US-08-971-395-4
7	25.4	21.2	9168	4	US-09-687-731-11
8	25	20.8	2266	1	US-08-453-472-1
9	25	20.8	2266	1	US-08-453-952-1
10	25	20.8	2266	2	US-08-484-993B-42
11	25	20.8	2266	2	US-08-862-903-1
12	25	20.8	2266	2	US-08-484-158B-42
13	25	20.8	2266	2	US-08-484-596A-42
14	25	20.8	2266	2	US-08-480-150A-42
15	25	20.8	2266	3	US-08-458-731-42
16	25	20.8	2266	3	US-08-149-223A-42
17	24.8	20.7	1230025	4	US-09-199-452A-1
18	24.6	20.5	1266	4	US-09-252-991A-1214
19	24.6	20.5	2262	4	US-09-252-991A-1353
20	24.6	20.5	2418	4	US-09-252-991A-1251
21	24.6	20.5	3910	4	US-09-120-653D-1
22	24.6	20.5	4403765	3	US-08-103-840A-2
23	24.6	20.5	4411529	3	US-09-103-840A-1
24	24.4	20.3	900	4	US-09-589-927-1
25	24.4	20.3	900	4	US-09-277-665-1
26	24.4	20.3	900	4	US-09-589-987-1
27	24.4	20.3	1077	4	US-09-252-991A-12157

C	28	24.4	20.3	1835	3	US-09-276-531-26	Sequence 26, Appl
	29	24.4	20.3	4103	4	US-09-620-312D-390	Sequence 390, App
	30	24.4	20.3	6674	4	US-09-620-312D-110	Sequence 110, App
C	31	24.2	20.2	716	2	US-08-211-718-1	Sequence 1, Appli
C	32	24.2	20.2	1608	2	US-08-211-718-8	Sequence 8, Appli
C	33	24.2	20.2	2257	1	US-08-452-722-6	Sequence 6, Appli
C	34	24.2	20.2	2257	1	US-08-404-731A-6	Sequence 6, Appli
C	35	24.2	20.2	2257	1	US-08-344-227-6	Sequence 6, Appli
C	36	24.2	20.2	2257	2	US-08-503-226B-6	Sequence 6, Appli
C	37	24.2	20.2	2257	3	US-08-721-458B-6	Sequence 6, Appli
C	38	24	20.0	2115	1	US-08-285-641-19	Sequence 19, Appl
C	39	24	20.0	2418	4	US-09-388-743-25	Sequence 25, Appl
C	40	23.6	19.7	1602	4	US-09-107-532A-1590	Sequence 1590, Ap
C	41	23.6	19.7	1858	3	US-08-742-185-96	Sequence 96, Appl
C	42	23.6	19.7	2169	1	US-08-379-496-1	Sequence 1, Appli
C	43	23.6	19.7	15239	1	US-08-390-878-17	Sequence 17, Appl
C	44	23.6	19.7	21119	4	US-09-453-702B-111	Sequence 111, App
C	45	23.6	19.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-846-762-1
; Sequence 1, Application US/08846762A
; Patent No. 5994072
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Burrows, Lori
; APPLICANT: Charter, Deborah
; APPLICANT: de Kievit, Teresa
; TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
; TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
; FILE REFERENCE: 6580-089
; CURRENT APPLICATION NUMBER: US/08/846.762A
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-1

Query Match	22.7%	Score 27.2;	DB 2;	Length 24417;
Best Local Similarity	51.1%;	Pred. No. 3.1;		
Matches	44;	Conservative	0;	Mismatches 28; Indels 0; Gaps 0;
Qy	37	GCACGGACTGCTGCGACGAGAAAACCTCTTACCGAAGTCTTCTATACCCAGGCTCAATA	96	
Db	14511	GCACGGACTAGCCTGGGAAGACCATTCAGACGATGCTCAGCTTTACGGAGCGGTCAA	14570	

Qy	97	GCCTCTCAAGGA	108	
Db	14571	GCCTCTCTCGGA	14582	

RESULT 2
US-08-229-287-3
; Sequence 3, Application US/08229287
; Patent No. 5530193
; GENERAL INFORMATION:
; APPLICANT: Clark Jr., John M.
; APPLICANT: Jilka, Joseph M.
; APPLICANT: Murry, Lynn E.
; APPLICANT: Scarafia, Liliana E.
; TITLE OF INVENTION: VIRUS RESISTANT CORN PLANTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc.
; STREET: 975 California Avenue
; CITY: Palo Alto


```
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5543 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 3' END OF INTRON 19 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
; US-08-687-080-101

Query Match          21.8%; Score 26.2; DB 2; Length 5543;
Best Local Similarity 60.8%; Pred. No. 4;
Matches 43; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 9 AACAGGCAAGGTAATGAAATGCTGAGGACGAGTGTGCTGCGAAGCAAACTCCCTTA 68
    |||||
Db 3056 AAAAGGCATTTATTAGCAATTATTGAGTCAGGCACTATCTGATTACATACATTC 3115
    |||||

QY 69 CCGAAGCTCTTC 79
    |||||
Db 3116 CTTTAATCTTC 3126

RESULT 5
US-09-385-982-326
; Sequence 326, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
```

```
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 326
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(657)
; OTHER INFORMATION: n = A,T,C or G
; US-09-385-982-326

Query Match          21.2%; Score 25.4; DB 3; Length 657;
Best Local Similarity 64.8%; Pred. No. 3.3;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 41 GGAATGCTGCGAAGCAAACTCCTTACCGAAGTCTTCTATACCCAGGCTCAA 94
    |||||
Db 335 GGAATGCAATTGGCATGAACNCAGCTTACTGNAGNCTTCCAAACCTGGGCTCAA 388
    |||||

RESULT 6
US-08-971-395-4
; Sequence 4, Application US/08971395
; Patent No. 6359197
; GENERAL INFORMATION:
; APPLICANT: Amasino, Richard M
; APPLICANT: No. 6359197, Yoo-Sun
; APPLICANT: Gan, Susheng
; TITLE OF INVENTION: Transgenic Plants with Altered
; TITLE OF INVENTION: Senescence Characteristics
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,395
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.94908
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-971-395-4

Query Match          21.2%; Score 25.4; DB 4; Length 5303;
Best Local Similarity 54.9%; Pred. No. 7.9;
Matches 50; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 16 AAGTAATGAAATGCTGAGGACGAGTGTGCTGCGAAGCAAACTCCTTACCGAAGT 75
    |||||
Db 4890 AAATATCAAGGATAAGAGGATTATGTTGCTTCCATGAGGCTTCTTATCCCAACT 4949
    |||||
```

QY 76 CTTCTATACCCAGGCTCAATACCGCTCAAG 106
Db 4950 ATATGAAACCCGGTTCAATACCCGGTCAAG 4980

RESULT 7

US-09-687-731-11
; Sequence 11, Application US/09687731
; Patent No. 6504080
; GENERAL INFORMATION:
; APPLICANT: Van Der Putten, Petrus Herman Maria
; TITLE OF INVENTION: Transgenic animal model for neurodegenerative disorders
; FILE REFERENCE: 'alpha-synuclein transgenic animals
; CURRENT APPLICATION NUMBER: US/09/687,731
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: GB 9924513.6
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 9168
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:DNA-construct
US-09-687-731-11

Query Match 21.2%; Score 25.4; DB 4; Length 9168;
Best Local Similarity 53.5%; Pred. No. 10; Mismatches 46; Indels 0; Gaps 0;
Matches 53; Conservative 0;
QY 5 CAGAACAGGCAAGGTAAATGAAATGCGCTGAGGACGAGCTGTGTCGGAACGAAACTC 64
Db 6447 CAGTAGCATGGAACATGGAACCAACCAATCCATTCCCTTTGCTGATATAACAGGCTC 6506
QY 65 CTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTC 103
Db 6507 CAAAGCAAAACCTGTCTACTGAGGCTCAAGAGCAGATC 6545

RESULT 8

US-08-453-472-1
; Sequence 1, Application US/08453472
; Patent No. 5626846
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,472
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2266

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: human

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

FEATURE:

NAME/KEY: ZP2

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: human ZP2 cDNA

US-08-453-472-1

Query Match 20.8%; Score 25; DB 1; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 6 AGAACAGGCAAGGTAAATGAAATGCTGAGGACGAGCTGTGTCGGAACGAAACTCC 65
Db 918 AGAAGCAACAAATGGCATGAAATGCAATTCAGCAAAACTCTGCTCAAAACGAAATTATC 977
QY 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAAAATGCTTACTCCATCATGTTACTTAGCTTCATCAAGCTGACCTTTC 1030

RESULT 9

US-08-453-952-1
; Sequence 1, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-May-1995
; CLASSIFICATION: 424

```

/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/038,948
/ FILING DATE: 26-MAR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/930,462
/ FILING DATE: 20-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/364,379
/ FILING DATE: 12-JUN-1989
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DOROTHY R. AUTH
/ REGISTRATION NUMBER: 36,434
/ REFERENCE/DOCKET NUMBER: 2026-4032 US4
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2266
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
/ ORIGINAL SOURCE:
/ ORGANISM: human
/ STRAIN:
/ INDIVIDUAL ISOLATE:
/ DEVELOPMENTAL STAGE:
/ HAPLOTYPE:
/ TISSUE TYPE:
/ CELL TYPE:
/ CELL LINE:
/ ORGANELLE:
/ FEATURE:
/ NAME/KEY: ZP2
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: human ZP2 cDNA
/
US-08-453-952-1

```

```

Query Match 20.8%; Score 25; DB 1; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 6 AGAAGCAGGCAAGTAATGAAATGCTGAGGCAAGGACTGTGCTGCGAAGCAAAACATCC 65
Db 918 AGAAGCAACAATGGCATGAATTCATTTGACGAAACTCTGCTCAAAACGAAATTTATC 977

QY 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAATAATGCTACTCCATCAGTTCTACTTAGCTTCACTCAAGCTGACCTTTC 1030

```

```

RESULT 10
US-08-484-993B-42
/ Sequence 42, Application US/08484993B
/ Patent No. 5837497
/ GENERAL INFORMATION:
/ APPLICANT: Harris Ph.D., Jeffrey D.
/ APPLICANT: Hsu, Kuang T.
/ APPLICANT: Podolski, Joseph S.
/ TITLE OF INVENTION: Materials and Methods for Immunococontraception
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 South Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk

```

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25.
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/484,993B
/ FILING DATE: 09-NOV-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/012,990
/ FILING DATE: 29-JAN-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/973,341
/ FILING DATE: 09-NOV-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clough, David W.
/ REGISTRATION NUMBER: 36,107
/ REFERENCE/DOCKET NUMBER: 31745
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6653
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2266 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2235
/
US-08-484-993B-42

```

```

Query Match 20.8%; Score 25; DB 2; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 6 AGAAGCAGGCAAGTAATGAAATGCTGAGGCAAGGACTGTGCTGCGAAGCAAAACATCC 65
Db 918 AGAAGCAACAATGGCATGAATTCATTTGACGAAACTCTGCTCAAAACGAAATTTATC 977

QY 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAATAATGCTACTCCATCAGTTCTACTTAGCTTCACTCAAGCTGACCTTTC 1030

```

```

RESULT 11
US-08-862-903-1
/ Sequence 1, Application US/08862903
/ Patent No. 5916768
/ GENERAL INFORMATION:
/ APPLICANT: DEAN, JURRIEN
/ TITLE OF INVENTION: CONTRACEPTIVE VACCINE
/ TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORGAN & FINNEGAN
/ STREET: 345 PARK AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 10154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/862,903
/ FILING DATE: 30-May-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-6800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2266
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cdna
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP2
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP2 cdna
US-08-862-903-1

Query Match 20.8%; Score 25; DB 2; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 6 AGAAGCAGGCAAGTAAATGAAATGCTGAGGCACGACTGTGCTGCGAAGCAAACTCC 65
Db 918 AGAAGCAACAATGGCATGAATTGCTTCAGCAAACTCTGCTCAAAACGAAATTATC 977

Qy 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAATAATGCTACTCCATCAGTTCTACTTAGCTTCACCTCAAGCTGACCTTTC 1030

RESULT 12
US-08-484-158B-42
; Sequence 42, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunoreception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 2266 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1...2235
US-08-484-158B-42

Query Match 20.8%; Score 25; DB 2; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 6 AGAAGCAGGCAAGTAAATGAAATGCTGAGGCACGACTGTGCTGCGAAGCAAACTCC 65
Db 918 AGAAGCAACAATGGCATGAATTGCTTCAGCAAACTCTGCTCAAAACGAAATTATC 977

Qy 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAATAATGCTACTCCATCAGTTCTACTTAGCTTCACCTCAAGCTGACCTTTC 1030

RESULT 13
US-08-484-596A-42
; Sequence 42, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunoreception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```
/ APPLICATION NUMBER: US/08/484,596A
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/149,223
/ FILING DATE: 09-NOV-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clough, David W.
/ REGISTRATION NUMBER: 36,107
/ REFERENCE/DOCKET NUMBER: 31745
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6653
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2266 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2235
/ US-08-484-596A-42

Query Match 20.8%; Score 25; DB 2; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 6 AGAAGCAGGCAAGTAAATGCTGAGGACGAGTGTGCTCGAAGCAAACTCC 65
Db 918 AGAAGCAAAATGGCATGAAATTCATTTGCAAGCAAACTGCTCAAAACGAAATTATC 977
QY 66 TTACCGAAGTCTTTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATC 118
Db 978 TGAATAATGCCTACTCCATCAGTTCTACTTAGTTCACTCAAGCTGACCTTTC 1030

RESULT 15
US-08-458-731-42
/ Sequence 42, Application US/08458731
/ Patent No. 6001599
/ GENERAL INFORMATION:
/ APPLICANT: Harris Ph.D., Jeffrey D.
/ APPLICANT: Hsu, Kuang T.
/ APPLICANT: Podolski, Joseph S.
/ TITLE OF INVENTION: Materials and Methods for Immunocontraception
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 South Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/458,731
/ FILING DATE: 09-NOV-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/012,990
/ FILING DATE: 29-JAN-1993
/ APPLICATION NUMBER: 07/973,341
/ FILING DATE: 09-NOV-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clough, David W.
/ REGISTRATION NUMBER: 36,107
/ REFERENCE/DOCKET NUMBER: 31745
/ TELECOMMUNICATION INFORMATION:
```

```

; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2235
; US-08-458-731-42

Query Match 20.8%; Score 25; DB 3; Length 2266;
Best Local Similarity 51.3%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 6 AGAAGCAGGCAAGGTAATGAAATGCTGAGGCGAGGACTGTGCTGCGAAGCAAACTCC 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 AGAAGCAGCAAAATGGCATGAAATTGCATTTACGCAAACTGTCTCAAAACGAAATTATC 977
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 66 TTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATC 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 TGAAAAATGCCTACTCCATCAGTTCTACTTACTTCACTCAAGCTGACCTTTC 1030
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 15, 2003, 08:09:00
Job time : 26.691 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 67.9536 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-74

Perfect score: 120

Sequence: 1 CGGTCAGAACAGCGAAGGT.....CTCAAGGAGAGAGCTATCAT 120

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata1/pubpna/PTCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata1/pubpna/PTCT_PUBCOMB.seq.*
7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	120	10	US-09-928-457-73
2	120	100.0	120	10	US-09-928-457-74
3	28	23.3	2598	10	US-09-938-842A-1646
4	27.6	23.0	261	9	US-09-923-876-5811
5	27.2	22.7	24417	14	US-10-216-209-1
6	27	22.5	525	12	US-10-027-632-51299
7	27	22.5	525	13	US-10-027-632-51299
8	26.6	22.2	394	10	US-09-960-352-13696
9	26.2	21.8	631	12	US-10-027-632-54625
10	26.2	21.8	631	13	US-10-027-632-54625
11	25.8	21.5	676	12	US-10-027-632-244583
12	25.8	21.5	676	13	US-10-027-632-244583
13	25.8	21.5	676	13	US-10-027-632-244583
14	25.8	21.5	676	13	US-10-027-632-244584
15	25.6	21.3	876	12	US-10-027-632-8784
16	25.6	21.3	876	12	US-10-027-632-8785

17	25.6	21.3	876	13	US-10-027-632-8784	Sequence 8784, Ap
18	25.6	21.3	876	13	US-10-027-632-8785	Sequence 8785, Ap
19	25.4	21.2	263	10	US-09-878-574-15498	Sequence 15498, A
20	25.4	21.2	565	12	US-10-027-632-234053	Sequence 234053, A
21	25.4	21.2	565	12	US-10-027-632-234054	Sequence 234054, A
22	25.4	21.2	565	13	US-10-027-632-234053	Sequence 234053, A
23	25.4	21.2	565	13	US-10-027-632-234054	Sequence 234054, A
24	25.4	21.2	657	11	US-09-871-161-326	Sequence 326, App
25	25.4	21.2	1027	11	US-09-822-846-69	Sequence 69, Appl
26	25.4	21.2	2121	12	US-10-027-632-250460	Sequence 250460, A
27	25.4	21.2	2121	12	US-10-027-632-250461	Sequence 250461, A
28	25.4	21.2	2121	13	US-10-027-632-250460	Sequence 250460, A
29	25.4	21.2	2121	13	US-10-027-632-250461	Sequence 250461, A
30	25.4	21.2	3378	10	US-09-917-800A-1443	Sequence 1443, Ap
31	25.2	21.0	565	12	US-10-029-386-10901	Sequence 10901, A
32	25.2	21.0	565	12	US-10-029-386-10901	Sequence 10901, A
33	25.2	21.0	565	12	US-10-027-632-191607	Sequence 191607, A
34	25.2	21.0	565	12	US-10-027-632-191608	Sequence 191608, A
35	25.2	21.0	565	13	US-10-027-632-191607	Sequence 191607, A
36	25.2	21.0	591	9	US-09-864-761-7709	Sequence 7709, Ap
37	25.2	21.0	987	14	US-10-156-761-198	Sequence 198, App
38	25.2	21.0	9025608	14	US-10-156-761-1	Sequence 1, Appl
39	25	20.8	556	12	US-10-029-386-12683	Sequence 12683, A
40	25	20.8	582	12	US-10-027-632-88852	Sequence 88852, A
41	25	20.8	582	12	US-10-027-632-88852	Sequence 88852, A
42	25	20.8	582	13	US-10-027-632-88853	Sequence 88853, A
43	25	20.8	582	13	US-10-027-632-88853	Sequence 88853, A
44	25	20.8	2434	14	US-10-198-846-11032	Sequence 11032, A
45	25	20.8	3732	11	US-09-764-891-5825	Sequence 5825, Ap

ALIGNMENTS

RESULT 1

US-09-928-457-73
; Sequence 73, Application US/09928457
; Patent No. US2002016403A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (OBB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-928-457-73

Query Match 100.0%; Score 120; DB 10; Length 120;
Best Local Similarity 100.0%; Pred. No. 2.9e-36;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGTCAGAACAGCGAAGGTATGAAAAATCGCTCAGGACGAGCTGTGCTCGGACGAAA 60
Db 1 CGGTCAGAACAGCGAAGGTATGAAAAATCGCTCAGGACGAGCTGTGCTCGGACGAAA 60

QY 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120
Db 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

RESULT 2

US-09-928-457-74
; Sequence 74, Application US/09928457
; Patent No. US20020164603A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30 (OEB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-928-457-74

Query Match 100.0%; Score 120; DB 10; Length 120;
Best Local Similarity 100.0%; Pred. No. 2.9e-36;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGAGAACAGGCAAGTATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAA 60
Db 1 CGGTGAGAACAGGCAAGTATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAA 60
QY 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120
Db 61 ACTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCGCTCAAGGAGAGAGCTATCAT 120

RESULT 3

US-09-938-842A-1646/c
; Sequence 1646, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1646

; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1646

Query Match 23.3%; Score 28; DB 10; Length 2598;
Best Local Similarity 60.5%; Pred. No. 3.2;
Matches 46; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 10 ACAGGCAAGTAATGAAATGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTAC 69
Db 782 ACTGGAAGAGTACGTAATCTCAGAGGAAGGACCGCTATGAAGGAAAAACCAAGC 723
QY 70 CGAAGTCTTCTATACC 85
Db 722 TTACTCTATTAACC 707

RESULT 4

US-09-923-876-5811/c
; Sequence 5811, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5811
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457772H1
; NAME/KEY: unsure
; LOCATION: 17, 70, 114, 131, 186-211, 249
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5811

Query Match 23.0%; Score 27.6; DB 9; Length 261;
Best Local Similarity 70.8%; Pred. No. 2;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 29 TGCTGAGGACGAGCTGTCTGCGAACGAAAACCTCTTACCGAGCTTTC 79
Db 175 TGCAGCGGCTCGGCTCTGGCCCGCAGCAACACCTTGGCGGANGCGTTC 125

RESULT 5

US-10-216-209-1
; Sequence 1, Application US/10216209
; Publication No. US20030124634A1
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Burrows, Lori
; APPLICANT: Charter, Deborah
; APPLICANT: De Kievit, Teresa De
; TITLE OF INVENTION: No. US20030124634A1e1 Proteins Involved in the Synthesis and Assen
; TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
; FILE REFERENCE: 6580-167
; CURRENT APPLICATION NUMBER: US/10/216,209
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/352,994
; PRIOR FILING DATE: 2001-05-29


```

; PRIOR APPLICATION NUMBER: US 08/846,762
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-216-209-1

Query Match      22.7%; Score 27.2; DB 14; Length 24417;
Best Local Similarity 61.1%; Pred. No. 14;
Matches 44; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 37 GCACGACTGTCTCGGACGAAACTCTTACCCAAAGTCTTCTATACCCAGGCTCAATA 96
Db 14511 GCACGACTGTCTCGGACGAAACTCTTACCCAAAGTCTTCTATACCCAGGCTCAATA 96

QY 97 GCGGCTCAAGGA 108
Db 14571 GCGTCTCTCGGA 14582

RESULT 6
US-10-027-632-51299
; Sequence 51299, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51299
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-51299

Query Match      22.5%; Score 27; DB 12; Length 525;
Best Local Similarity 53.5%; Pred. No. 4.4;
Matches 54; Conservative 1; Mismatches 46; Indels 0; Gaps 0;

QY 2 GGTGAGAAACAGCGAGGTAAATGAAATGCTTGAGGCACGAGTGTGCTGCGAAGCAAAA 61
Db 39 GGCAGGCTGGAGGAGCCACTGACATTGCAAGAGGCACAGATAATGCTCTTTGGACAA 98

QY 62 CTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCT 102
Db 99 CTGCCAGAGTGTGTCAAAYCAGCGGGGCCAAGGGGCGCTCT 139

RESULT 7
US-10-027-632-51299
; Sequence 51299, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51299
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-51299

Query Match      22.5%; Score 27; DB 12; Length 525;
Best Local Similarity 53.5%; Pred. No. 4.4;
Matches 54; Conservative 1; Mismatches 46; Indels 0; Gaps 0;

QY 2 GGTGAGAAACAGCGAGGTAAATGAAATGCTTGAGGCACGAGTGTGCTGCGAAGCAAAA 61
Db 39 GGCAGGCTGGAGGAGCCACTGACATTGCAAGAGGCACAGATAATGCTCTTTGGACAA 98

QY 62 CTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCT 102
Db 99 CTGCCAGAGTGTGTCAAAYCAGCGGGGCCAAGGGGCGCTCT 139

RESULT 8
US-09-960-352-13696
; Sequence 13696, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13696
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 58-LIB34-079-Q1-E1-G10
US-09-960-352-13696

Query Match      22.2%; Score 26.6; DB 10; Length 394;
Best Local Similarity 52.2%; Pred. No. 5.6;
Matches 59; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 8 AACAGCGCAAGGTAAATGAAATGCTTGAGGCACGAGTGTGCTGCGAAGCAAACTCCTT 67

```

```
Db 128 AAACCGTCAAGATTGTAATGTTGACCTTTGGCAAGAAATGGGATTCATATGAACAAGGATT 187
QY 68 ACCGAGTCTTCTATACCCAGGCTCAATAGCCGCTCAAGGAGAGAGCTATCAT 120
Db 188 TGGGAATATTGAACACCAATGATGAGGAAGAAATACGTTGGTGTAAAT 240

RESULT 9
US-10-027-632-54625
; Sequence 54625, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54625
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54625

Query Match 21.8%; Score 26.2; DB 12; Length 631;
Best Local Similarity 56.3%; Pred. No. 9.5;
Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 25 AAAATGCTGAGCGGAGTGTGCTGGCAAGAACTCCTTACCGAAGTCTTCTATAC 84
Db 357 AAAATACCTTAGACTGGGTAATTGATAAGAACAGAAATTTATTTCTGGAGGCTGGGAGT 416

QY 85 CCAGGCTCAATAGCCGCTCAAGGAGAG 111
Db 417 CCAAGATCAAAAGCACTGCGAGGTGTG 443

RESULT 10
US-10-027-632-54625
; Sequence 54625, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54625
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54625

Query Match 21.8%; Score 26.2; DB 12; Length 631;
Best Local Similarity 56.3%; Pred. No. 9.5;
Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 25 AAAATGCTGAGCGGAGTGTGCTGGCAAGAACTCCTTACCGAAGTCTTCTATAC 84
Db 357 AAAATACCTTAGACTGGGTAATTGATAAGAACAGAAATTTATTTCTGGAGGCTGGGAGT 416

QY 85 CCAGGCTCAATAGCCGCTCAAGGAGAG 111
Db 417 CCAAGATCAAAAGCACTGCGAGGTGTG 443

RESULT 11
US-10-027-632-244583/c
; Sequence 244583, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244583
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-244583

Query Match 21.5%; Score 25.8; DB 12; Length 676;
Best Local Similarity 53.5%; Pred. No. 14;
Matches 54; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 2 GGTGAGAAACAGGCAAGGTAATGAAATAATCCCTGAGGCGGAGTGTGCTGCGAAGCAAAA 61
Db 258 GGGCAGGCTGGAGGAGCCACTGACATTCGAAGAGGCAGAGATAATGCTGCTTTGGAACA 199

QY 62 CTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCT 102
Db 198 CTGCCAGAGTGTGTGACACCAACCGCGGGCCAAAGGGGCTCT 158

RESULT 12
US-10-027-632-244584/c
```

```
/ Sequence 244584, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 244584
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-244584

Query Match 21.5%; Score 25.8; DB 12; Length 676;
Best Local Similarity 53.5%; Pred. No. 14;
Matches 54; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 2 GGTGAGAAACAGCAGGTAAATGAAATGCTGAGGCACGACTGTGTCGGAACGAAAA 61
Db 258 GGCAGGCTGGAGGAGCCACTGACATTGCAAGAGGCACAGATAATGCTGTTGGACAA 199

QY 62 CTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCT 102
Db 198 CTGCCAGAGTGTGCAACAGCCGGGGCCCAAGGGGCTCT 158

RESULT 14
US-10-027-632-244584/c
/ Sequence 244584, Application US/10027632
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 244584
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-244584

Query Match 21.5%; Score 25.8; DB 13; Length 676;
Best Local Similarity 53.5%; Pred. No. 14;
Matches 54; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 2 GGTGAGAAACAGCAGGTAAATGAAATGCTGAGGCACGACTGTGTCGGAACGAAAA 61
Db 258 GGCAGGCTGGAGGAGCCACTGACATTGCAAGAGGCACAGATAATGCTGTTGGACAA 199

QY 62 CTCCTTACCGAAGTCTTCTATACCCAGGCTCAATAGCCGCT 102
Db 198 CTGCCAGAGTGTGCAACAGCCGGGGCCCAAGGGGCTCT 158

RESULT 15
US-10-027-632-8784
/ Sequence 8784, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8784
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8784

Query Match      21.3%; Score 25.6; DB 12; Length 876;
Best Local Similarity 57.5%; Pred. No. 18;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY      8 AACACGGCAAGGTATGAAATGCTGAGGCGGACTGTGCTGCGAAGCAACTCGTT 67
Db      446 AAACCTTGCAAGAGACAGAAAAGGGCTTACGTATGGCAATTATTAGATAAAACTATTT 505

QY      68 ACCGAAGTCTTCTATACCCA 87
Db      506 TCTAGGATAATATACCCTA 525

Search completed: November 15, 2003, 08:31:54
Job time : 74.0536 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 1190.78 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-77

Perfect score: 269
Sequence: 1 CGAGCATAAATCGTTATT.....GCAAGGTTATTATCAACG 269

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.ov.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.tod.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	269	100.0	269	1	AF169451	AF169451 Neisseria
2	269	100.0	269	6	A68906	A68906 Sequence 77
3	269	100.0	269	6	BD063006	BD063006 DNA and s
4	267.4	99.4	349061	1	NMA222491	AL162753 Neisseria
5	264.2	98.2	858	6	AX024058	AX024058 Sequence
6	246.6	91.7	858	6	AX024111	AX024111 Sequence
7	121.8	45.3	270	6	A68917	A68917 Sequence 88
8	121.8	45.3	270	6	BD063017	BD063017 DNA and s
9	121.8	45.3	271	1	AF169464	AF169464 Neisseria
10	110.8	41.2	270	1	AF169428	AF169428 Neisseria
11	67.8	25.2	4884	1	HGDIIM	X55141 H.giganteus
12	42.6	15.8	164490	2	AC079077	AC079077 Homo sapi
13	42.2	15.7	63215	2	AC061968	AC061968 Homo sapi
14	41	15.2	185539	2	AC018349	AC018349 Homo sapi
15	39.4	14.6	125746	9	AC018279	AC018279 Homo sapi
16	39.4	14.6	142715	9	AC008387	AC008387 Homo sapi
17	38.8	14.4	254912	2	AC123482	AC123482 Rattus no
18	38.8	14.4	333352	2	AC113855	AC113855 Rattus no
19	38.8	14.4	333352	2	AC125577	AC125577 Rattus no
20	38.2	14.2	123098	9	AC008799	AC008799 Homo sapi
21	38.2	14.2	151031	9	AC008923	AC008923 Homo sapi
22	38	14.1	187166	2	AC138297	AC138297 Mus muscu
23	37.6	14.0	110000	2	AC117022_1	Continuation (2 of
24	37.6	14.0	121922	2	AP000576	AP000576 Homo sapi
25	37.6	14.0	161582	9	AP001636	AP001636 Homo sapi
26	37.6	14.0	161586	9	AC090309	AC090309 Homo sapi
27	37.6	14.0	169670	9	AC023908	AC023908 Homo sapi
28	37.6	14.0	239844	2	AC134274	AC134274 Rattus no
29	37.4	13.9	144201	10	AC087218	AC087218 Rattus no
30	37.4	13.9	227796	2	AC131418	AC131418 Rattus no
31	37.4	13.9	248922	2	AC129700	AC129700 Rattus no
32	37.2	13.8	44298	3	U58751	U58751 Caenorhabdi
33	37.2	13.8	143092	3	CEV17G7B	AL023828 Caenorhab
34	37	13.8	192969	10	AC122373	AC122373 Mus muscu
35	36.8	13.7	2578	3	AY051762	AY051762 Drosophil
36	36.8	13.7	3170	3	AB049434	AB049434 Drosophil
37	36.8	13.7	20175	2	AC019990	AC019990 Drosophil
38	36.8	13.7	55307	2	AC014684	AC014684 Drosophil
39	36.8	13.7	150391	2	AC011764	AC011764 Drosophil
40	36.8	13.7	162345	3	AC006415	AC006415 Drosophil
41	36.8	13.7	170498	3	AC008094	AC008094 Drosophil
42	36.8	13.7	186695	3	AC013351	AC013351 Drosophil
43	36.8	13.7	244757	3	AE003678	AE003678 Drosophil
44	36.8	13.7	298116	3	AE003782	AE003782 Drosophil
45	36.6	13.6	179607	2	AC102778	AC102778 Mus muscu

ALIGNMENTS

RESULT 1
AF169451
LOCUS AF169451
DEFINITION Neisseria meningitidis strain Z2491 clone Cm040 unknown sequence.
ACCESSION AF169451
VERSION AF169451.1 GI:9754659
KEYWORDS Neisseria meningitidis
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 269)
Perrin A., Nassif, X. and Tinsley, C.R.
AUTHORS Identification of regions of the chromosome of Neisseria
TITLE meningitidis and Neisseria gonorrhoeae which are specific to

```

pathogenic Neisseriae
Unpublished
REFERENCE 2 (bases 1 to 269)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        Location/Qualifiers
            1..269
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /strain="Z2491"
            /db_xref="taxon:487"
            /clone="Cm040"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
    Query Match 100.0%; Score 269; DB 1; Length 269;
    Best Local Similarity 100.0%; Pred. No. 9e-60;
    Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
Db 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
QY 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
Db 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
QY 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
Db 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
QY 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
Db 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
Db 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269

RESULT 3
BD063006 269 bp DNA linear PAT 27-AUG-2002
LOCUS DNA and specific proteins or peptides of the Neisseria meningitidis
DEFINITION species bacteria, method for obtaining them and their biological
applications.
ACCESSION BD063006
VERSION BD063006.1 GI:22608609
KEYWORDS JP 2001504684-A/68.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.I., Vinals,C. and
Marker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 68 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/68
PD 10-APR-2001
PF 11-JUL-1997 JP 1998505895
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC Cl2N15/31.C07K14/22.C07K16/12.A61K39/095.C12Q1/68.G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..269
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
    Query Match 100.0%; Score 269; DB 6; Length 269;
    Best Local Similarity 100.0%; Pred. No. 9e-60;
    Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
Db 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
QY 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
Db 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
QY 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
Db 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
QY 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
Db 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
Db 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269

RESULT 2
A68906 269 bp DNA linear PAT 06-MAY-1999
LOCUS Sequence 77 from Patent WO9802547.
DEFINITION A68906
ACCESSION A68906
VERSION A68906.1 GI:4759825
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Marker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
JOURNAL Patent: WO 9802547-A 77 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
Other publication FR 2751000 19980116.
COMMENT Location/Qualifiers
FEATURES
    source
        1..269
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
    Query Match 100.0%; Score 269; DB 6; Length 269;
    Best Local Similarity 100.0%; Pred. No. 9e-60;
    Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
Db 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
QY 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
Db 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
QY 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
Db 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
QY 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
Db 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
Db 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269

RESULT 3
BD063006 269 bp DNA linear PAT 27-AUG-2002
LOCUS DNA and specific proteins or peptides of the Neisseria meningitidis
DEFINITION species bacteria, method for obtaining them and their biological
applications.
ACCESSION BD063006
VERSION BD063006.1 GI:22608609
KEYWORDS JP 2001504684-A/68.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.I., Vinals,C. and
Marker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 68 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/68
PD 10-APR-2001
PF 11-JUL-1997 JP 1998505895
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC Cl2N15/31.C07K14/22.C07K16/12.A61K39/095.C12Q1/68.G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..269
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
    Query Match 100.0%; Score 269; DB 6; Length 269;
    Best Local Similarity 100.0%; Pred. No. 9e-60;
    Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
Db 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATG 60
QY 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
Db 61 ATTTTATTGAGATCGGTTCGGAACAGAGGGAAGGAAACAAAGCTCCCGTCGGGA 120
QY 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
Db 121 GAATTCACACGGGTAAAAAGGCTTGGTAAATTTGGCATTATTCGGGCTTGGCAACAAA 180
QY 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
Db 181 TTGAAATTTCTACTATCCAGGGAAACGAAAGGTTACTTTTACTTTTGGATTATGACAGA 240
QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
Db 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269

RESULT 2
A68906 269 bp DNA linear PAT 06-MAY-1999
LOCUS Sequence 77 from Patent WO9802547.
DEFINITION A68906
ACCESSION A68906
VERSION A68906.1 GI:4759825
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Marker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
JOURNAL Patent: WO 9802547-A 77 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
Other publication FR 2751000 19980116.
COMMENT Location/Qualifiers
FEATURES
    source
        1..269
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
    Query Match 100.0%; Score 269; DB 6; Length 269;
    Best Local Similarity 100.0%; Pred. No. 9e-60;
    Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 121 GAATTCACCGGTAAAAAGGCTTGGTAAATGGCATTTATCGGCTTGGCAACAAA 180
Db 121 GAATTCACCGGTAAAAAGGCTTGGTAAATGGCATTTATCGGCTTGGCAACAAA 180

QY 181 TTGAATTTCTACTATCCAGGGAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 240
Db 181 TTGAATTTCTACTATCCAGGGAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 240

QY 241 TTCGAGAACCAAGGTTATTATCAACCG 269
Db 241 TTCGAGAACCAAGGTTATTATCAACCG 269

RESULT 4
NMA22491
LOCUS NMA22491 349061 bp DNA linear BCT 02-SEP-2002
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 2/7.
ACCESSION AL162753 AL157959
VERSION AL162753.2 GI:7379120
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 349061)
Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Felwell,T., Hamlin,N.,
Kholodny,S., Jagers,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skellton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
NATURE 404 (6777), 502-506 (2000)
JOURNAL
MEDLINE
2022556
PUBMED
10761919
REFERENCE
2 (bases 1 to 349061)
Parkhill,J.
DIRECT SUBMISSION
Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
NOTES
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/projects/N\_meningitidis/).
FEATURES
Location/Qualifiers
1..349061
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/note="serogroup: A"
complement(24..206)
/note="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
310.39"
/label=ATR
RBS 209..212
Gene 220..681
CDS 220..681
/gene="NMA0368"
/note="NMA0368", probable integral membrane protein, len:
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83669.1"
/db_xref="GI:7379121"
/translation="MQEQNRKSSFPFIVMLLVSVALLASLNVAFLGNHGSMEGLTV

```

```

LILGSIFASLDIRYCAVYVNWLAALVLLALRKKKVPVHAAFWGLALVAFSVKAVVY
DEAGNTSDIVRYGAGFVLYAAFAVASIGTFAGKKRKAASAADGTNDY"
671..1492
/gene="hemK"
671..1492
/gene="hemK"
/note="NMA0369, hemK, HemK protein, len: 273 aa; similar
to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
scores; E(): 0.42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
/codon_start=1
/transl_table=11
/product="HemK protein"
/protein_id="CAB83670.1"
/db_xref="GI:7379122"
/translation="MTDFKMLGSKLPKNEARMQLQVSEYRVOQLLRGGEEMPDEV
RQRADRLAQRRLNGEVPAYILGAREFYGRRTFVNPVSLIPREPTEHLVEAVLAKLPEN
GRVDLGTGGAVATVVALRPPDAFVRASDISPEALTARKNAADLGARVEFAYGSWF
DTMPSGKMDLIYVSNPPYIENGKHLISQSDLRPREPOLATDFSGLSCLRTLAQGAP
DRLAEGFLJLLEHGFQDGAAVRNGVLAENGSGVETLPDLAAGLDRLVTGKYNKHUK"
831..840
/gene="hemK"
/note="Core DNA uptake sequence: gcgctgtgaa"
/label=DUS
1162..1171
/gene="hemK"
/note="Core DNA uptake sequence: gcgctgtgaa"
/label=DUS
1187..1207
/gene="hemK"
/note="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/note="Core DNA uptake sequence: gcgctgtgaa"
/label=DUS
1578..1568
/gene="NMA0370"
1578..2966
/gene="NMA0370"
/note="NMA0370", probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325 HAEIN
P44640 hypothetical protein H10325 (450 aa), fasta scores;
E(): 0.49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83671.1"
/db_xref="GI:7379123"
/translation="MNAVVAIVMLVLSLSRVHVLSTIGAFVGAVAGMPLQNTA
DAAGVSOAGIIPVKNKLGSGAKIALSVAMLGAPAMATHTSGLPOLAGAVRKLNR
GMPDSVRSCEGVKVLKLLSILVGMNSQNIPIHTAFIPMIVPELLVFNRLKIDR
RLIACVITFLVTVTFPLPYGGAIFLINEILLGNHSAAPCLDKNVNNVAAKIPAL
GMIALLLAHVHYRKPRLYQSNADAGNAANRPQPSAYRSLAAVAIVCFAIQL
MYEDSLVIGALGFVFMGLGVNRDKANDVFGEIKMMAMVGFIMIAAQGFVAVNA
TGHQIVLVESSMAIFGNSKGAALAMLVGLLVMTGIGTSFSTLPIIAIYVPLCVGL
GFSPLATVITGTAGALGDAGSPASDSTLGTWGLNADGQHDHDSVITFTIHYNIP
LLIAGMTAAWVL"
2967..3191
/gene="slyX"
2967..3191
/gene="slyX"
/note="NMA0371, slyX, SLX protein homolog, len: 74 aa;
similar to SLXV_HAEIN P44759 SLXV protein homolog (73 aa),
fasta scores; E(): 0.0017.33.8% identity in 74 aa
overlap, and SLXV_ECOLI P30857 SLXV protein (72 aa), fasta
scores; E(): 0.24.32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=11
/product="SLXV protein homolog"
/protein_id="CAB83672.1"

```


Db 278 GAATTCACAGCGGTAAACAAAGCGCTTGGTAAATTGGCAATTATTTCGGCTTGGCAACAAAA 337
 QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGGTACTTTTACTTTTGGATTATGCAGAGA 240
 Db 338 TTGAATTTCTACTATCCAGGAAACGAAAGGGTACTTTTACTTTTGGATTATGCAGAGA 397
 QY 241 TTGAGAGACGAGGGTATTATCAACCG 269
 Db 398 TTGAGAGACGAGGGTATTATCAACCG 426

RESULT 6
 AX024111
 LOCUS AX024111 858 bp DNA linear PAT 15-SEP-2000
 DEFINITION Sequence 54 from Patent FR2785293.
 ACCESSION AX024111
 VERSION AX024111.1 GI:10184422
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria gonorrhoeae
 Neisseria gonorrhoeae
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 Neisseriaceae; Neisseria.

REFERENCE 1
 AUTHORS Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
 Bouchardon, A. and Renaud, M.G.
 JOURNAL Patent: FR 2785293-A 54 05-MAY-2000;
 PASTEUR MERIEUX SERUMS VACC (FR)

FEATURES
 source
 1..858
 /organism="Neisseria gonorrhoeae"
 /mol_type="genomic DNA"
 /db_xref="taxon:485"
 1..858
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC08931.1"
 /db_xref="GI:10184423"
 /translation="MSEELKMSFPTVIEHLGVKMYSHTVPAIAELIANAVDACATE
 VEVRLFDKPEKIVIKDNGIGMSFDEINDFYLRIGNRREKQASPCGRITPGKKGLG
 KIALFRLNKIISITQGNERTVFDIAEIKRSRIYQPEFKQSVKRENTGTTIT
 LTELKKQYPDNVYVGHSLRFPDPAODFKIKVSLNGSEPRIIDGNLKNLIVTPQE
 FTDSSSHF"

BASE COUNT 302 a 151 c 178 g 227 t
 ORIGIN

Query Match 91.7%; Score 246.6; DB 6; Length 858;
 Best Local Similarity 94.8%; Pred. No. 5.8e-54;
 Matches 255; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTTATAAGATAATGGTATAGGAACGAGCTTCGATGAAATCAATG 60
 Db 158 CGGAGCATAAATCGTTTATAAGATAATGGTATAGGAATGAGCTTCGATGAAATCAATG 217
 QY 61 ATTTTATTGAGATCGGTCGGAACAGAGGAGAGAAAAACAAGCTCCCGTCGGAA 120
 Db 218 ATTTTATTGAGATCGGTCGGAACAGAGGAGAGAAAAACAAGCTCCCGTCGGAA 277
 QY 121 GAATTCACAGGTAAGAACGCTTGGTAATTCGATTATTCGGCTTCGACAAAA 180
 Db 278 GAATTCACAGGTAAGAACGCTTGGTAATTCGATTATTCAGGCTTGGCAACAAA 337
 QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGGTACTTTTACTTTTGGATTATGCAGAGA 240
 Db 338 TCGAATCTCTACTATCCAGGAAACGAAAGGGTACTTTTACTTTTGGATTATGCAGAGA 397
 QY 241 TTGAGAGACGAGGGTATTATCAACCG 269
 Db 398 TTGAGAGAGTGAAGCGTATTATCAACCG 426

RESULT 7

A68917/c
 LOCUS A68917 270 bp DNA linear PAT 06-MAY-1999
 DEFINITION Sequence 88 from Patent WO9802547.
 ACCESSION A68917
 VERSION A68917.1 GI:4759836
 KEYWORDS
 SOURCE
 ORGANISM
 unidentified
 unclassified.
 1 (bases 1 to 270)
 AUTHORS Nassif, X., Tinsley, C., Achtman, M., Ruelle, J., Vinals, C. and
 Merker, P.
 TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
 SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
 APPLICATIONS
 JOURNAL Patent: WO 9802547-A 88 22-JAN-1998;
 INST NAT SANTE RECH MED (FR)
 COMMENT Other publication: FR 2751000 19980116.
 FEATURES
 source
 1..270
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 72 a 67 c 42 g 89 t
 ORIGIN

Query Match 45.3%; Score 121.8; DB 6; Length 270;
 Best Local Similarity 98.4%; Pred. No. 2.4e-21;
 Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTTATAAGATAATGGTATAGGAACGAGCTTCGATGAAATCAATG 60
 Db 125 CGGAGCATAAATCGTTTATAAGATAATGGTATAGGAATCAATTCGATGAAATCAATG 66
 QY 61 ATTTTATTGAGATCGGTCGGAACAGAGGAGAGAAAAACAAGCTCCCGTCGGAA 120
 Db 65 ATTTTATTGAGATCGGTCGGAACAGAGGAGAGAAAAACAAGCTCCCGTCGGAA 6
 QY 121 GAATTC 125
 Db 5 GAATTC 1

RESULT 8
 BD063017/c
 LOCUS BD063017 270 bp DNA linear PAT 27-AUG-2002
 DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis
 species bacteria, method for obtaining them and their biological
 applications.
 ACCESSION BD063017
 VERSION BD063017.1 GI:22608620
 KEYWORDS JP 2001504684-A/79.
 SOURCE
 ORGANISM
 unidentified
 unclassified.
 1 (bases 1 to 270)
 AUTHORS Nassif, X., Tinsley, C., Achtman, M., Ruelle, J.L., Vinals, C. and
 Merker, P.
 TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
 species bacteria, method for obtaining them and their biological
 applications
 JOURNAL Patent: JP 2001504684-A 79 10-APR-2001;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
 PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
 SMITHKLINE BEECHAM
 COMMENT PN JP 2001504684-A/79
 PD 10-APR-2001
 PF 11-JUL-1997 JP 1998505685
 PI 12-JUL-1996 FR 96/08768
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
 CARLA VINALS.
 PI PETEA MERKER
 PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC

Herpetosiphon giganteus strain Hpa2
 Gene 106 (1), 87-92 (1991)
 MEDLINE 92039068
 PUBMED 1937045
 COMMENT See also X55137-X55143
 FEATURES Location/Qualifiers
 1..4884
 /organism="Herpetosiphon aurantiacus"
 /mol_type="genomic DNA"
 /strain="Hpa2"
 /db_xref="taxon:65"
 884..1300
 /gene="orf 15"
 884..1300
 /gene="orf 15"
 /codon_start=1
 /transl_table=1
 /protein_id="CAA38940.1"
 /db_xref="GI:48772"
 /db_xref="SWISS-PROT:P25279"
 /translation="MWTAGYLTREAIGLCRNDVAHEAGTNNVQIMRIEKGIDTR
 GSLLSVRAVNFNAEHIQFLFMLEATEEDGRNLAIWINRETTISAIDEFADVKK
 NKVSEALKIQLEALDPSSLRLLGVQSLDRNR"
 1749..2813
 /gene="hgidiim"
 1749..2813
 /gene="hgidiim"
 /codon_start=1
 /transl_table=1
 /product="methyltransferase"
 /protein_id="CAA38941.1"
 /db_xref="GI:48773"
 /db_xref="SWISS-PROT:P25265"
 /translation="MVGAVIDLFCGVGLTHGILLEGFVLACIDNDPCKYAEQNN
 RTRPIKSEIVSENGRELNALPNKHILVGCAPQDFQYTKSKSTGKTKWLLTEFS
 RLRIEPEPDIQNGIGMSIEDANTKFLRIGHQKREDSANTISGRHYMGRKGIGILA
 DRVLFAKGVGKIKIPPTPTENRYTRVDVIGSLATNYSGGHWGDSWHAASRLIED
 NLRIIOHSPGVGTWDPSELIAECHKESGESYGVYGRMWDKVAPTITTCCKGYG
 NRGFHEQDRAQLSREALLQTFPSYQFAPEGQLKFTVSRQIGNAVFVALGRVIA
 KSTKRLGLEHQRVRILL"
 2782..4617
 /gene="orf 68"
 2782..4617
 /gene="orf 68"
 /codon_start=1
 /transl_table=1
 /protein_id="CAA38942.1"
 /db_xref="GI:48774"
 /db_xref="SWISS-PROT:P25280"
 /translation="MDSGYEISFLNTLNLHSEGLYSNLFYKLKLVANAWDADATE
 VSNIRQDEIVIQNGIGMSIEDANTKFLRIGHQKREDSANTISGRHYMGRKGIGILA
 IFGIANIAEYSCDGVPHGFIIHKGDIERGISDVLYRPSVPCEALSIBSGKII
 LREIKSIGNAEKTLDLARFTIINNNSFVLIINTPINDNRDYLKQVFLWIL
 GSESSKYANFTKLKSFETNLGDGSMGVKIGTVIRPSDIPNSQPYHIYFAHG
 KMLQEDILDI DAGRYROYIIGIEADFNDSDDDDIITSNEKRIKOTDPYLLKLE
 YKIDIMRVIASTWTKLREYSPNPKKEVNDSSLSKDSSEKENTWASSDSSTATE
 NASSDSSTWASSETNDGEVDNSFFDDDIPEPSPPKQEIITAFREMKLVQNSN
 IPNQMNILYDIQQAAYAKGTSFKACIVMLGAILEGVMGLCIQRTDVLVLIJALPO
 VKPLSDLPKPKADRVLAQYITPESFODCKEIELCVGNTKLGVDILQVRN
 SIHPGSVLKMDQLAFNFHQSAVGVIKHLINLVILNWPSP"
 1500 a 882 c 1055 g 1447 t

BASE COUNT 1500 a 882 c 1055 g 1447 t
 ORIGIN

Query Match 25.2%; Score 67.8; DB 1; Length 4884;
 Best Local Similarity 53.6%; Pred. No. 2.4e-07;
 Matches 141; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 7 ATAAATCGTTTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATGATTTT 66
 |||||
 Db 2936 ATCAATAGTCATCCAGGATAATGATAGGAATGTCAATTGAGGATGCTAATACTAAAT 2995
 |||||
 QY 67 ATTGAGATCGTTCGGAAGGAGGAAGAAACAAACCTCCCGCGGGAAGATTC 126
 |||||

2996 TCTTACGAATTGGCATCAAAAGAGAGAGAGATTACGTAACACTATAGTGCAGCGCATG 3055
 QY 127 CAACGGGTAATAAGCCCTTGGTAATTCGCAATTATTCGGGTTGSCAACAAAATGAAA 186
 |||||
 Db 3056 TTATGGCGAGGAAGGGATCGGTATCTGCAATATTTGGTATAGCCCAATATTGCTGAG 3115
 |||||
 QY 187 TTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGATTCGAA 246
 |||||
 Db 3116 TTTATCTTGTAGGATGAGTACCCCATGGATTTATATTCATAAGCGGATATAGAAA 3175
 |||||
 QY 247 GAAGCAAGGTTATTTATCAACCG 269
 |||||
 Db 3176 GAGGATAAGTAGCGATGTACG 3198
 |||||

RESULT 12
 AC079077
 LOCUS
 DEFINITION Homo sapiens chromosome 5 clone RP11-424A10 map 5, WORKING DRAFT
 SEQUENCE, 42 unordered pieces.
 AC079077
 AC079077.2 GI:11178110
 HTG; HTGS PHASE1; HTGS_DRAFT.
 VERSION
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 164490)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 5, clone RP11-424A10
 Unpublished
 2 (bases 1 to 164490)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
 Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
 Chepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
 FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
 Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
 Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
 MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
 O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
 Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
 Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
 Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Turrell,A., Travers,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (17-AUG-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164490)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
 Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
 Chepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
 FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
 Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
 Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
 MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
 O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
 Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
 Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
 Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zimmer, A., and Zody, M.

JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 15, 2000 this sequence version replaced gi:9838006.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L10780

Center clone name: 424_A_10

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 140058 bases at least Q40

Consensus quality: 152620 bases at least Q30

Consensus quality: 157604 bases at least Q20

Insert size: 183000; agarose-fp

Insert size: 160390; sum-of-contigs

Quality coverage: 2.7 in Q20 bases; agarose-fp

Quality coverage: 3.1 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 42 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 379: contig of 379 bp in length
380 479: gap of 100 bp
480 15053: contig of 14574 bp in length
15054 15153: gap of 100 bp
15154 15720: contig of 567 bp in length
15721 15820: gap of 100 bp
15821 17135: contig of 1315 bp in length
17136 17235: gap of 100 bp
17236 18343: contig of 1108 bp in length
18344 18443: gap of 100 bp
18444 19570: contig of 1127 bp in length
19571 19670: gap of 100 bp
19671 21870: contig of 2200 bp in length
21871 21970: gap of 100 bp
21971 23558: contig of 1588 bp in length
23559 23658: gap of 100 bp
23659 24734: contig of 1076 bp in length
24735 24834: gap of 100 bp
24835 26227: contig of 1393 bp in length
26228 26327: gap of 100 bp
26328 28306: contig of 1979 bp in length
28307 28405: gap of 100 bp
28407 29829: contig of 1423 bp in length
29830 29929: gap of 100 bp
29930 31365: contig of 1436 bp in length
31366 31465: gap of 100 bp
31466 33041: contig of 1576 bp in length
33042 33141: gap of 100 bp
33142 35033: contig of 1898 bp in length
35034 35133: gap of 100 bp
35134 37112: contig of 1973 bp in length
37113 38733: contig of 100 bp
38734 38833: gap of 100 bp
38834 41123: contig of 2290 bp in length

41124 41223: gap of 100 bp
41224 43464: contig of 2241 bp in length
43465 43564: gap of 100 bp
43565 46245: contig of 2681 bp in length
46246 46345: gap of 100 bp
46346 49174: contig of 2829 bp in length
49175 49274: gap of 100 bp
49275 51180: contig of 1916 bp in length
51181 51290: gap of 100 bp
51291 53683: contig of 2403 bp in length
53684 53783: gap of 100 bp
53784 57615: contig of 3822 bp in length
57616 57715: gap of 100 bp
57716 60918: contig of 3203 bp in length
60919 61018: gap of 100 bp
61019 63663: contig of 2645 bp in length
63664 63763: gap of 100 bp
63764 67175: contig of 3412 bp in length
67176 67275: gap of 100 bp
67276 70782: contig of 3507 bp in length
70783 75448: contig of 4566 bp in length
75449 75548: gap of 100 bp
75549 79599: contig of 4051 bp in length
79600 79699: gap of 100 bp
79700 84596: contig of 4897 bp in length
84597 84696: gap of 100 bp
84697 89528: contig of 4832 bp in length
89529 89628: gap of 100 bp
89629 94192: contig of 4564 bp in length
94193 94292: gap of 100 bp
94293 99758: contig of 5466 bp in length
99759 105474: gap of 100 bp
105475 105574: contig of 5616 bp in length
105575 111165: gap of 100 bp
111166 117392: contig of 6127 bp in length
117393 117492: gap of 100 bp
117493 128939: contig of 11447 bp in length
128940 129039: gap of 100 bp
129040 137760: contig of 8721 bp in length
137761 137860: gap of 100 bp
137861 149949: contig of 12089 bp in length
149950 150049: gap of 100 bp
150050 164298: contig of 14249 bp in length
164299 164398: gap of 100 bp
164399 164490: contig of 92 bp in length

FEATURES

source

1. 164490
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="s"
/clone="RP11-424A10"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. 379
/note="assembly_fragment"
clone_end.SP6
vector_side:left"

misc_feature

480..15053
/note="assembly_fragment"

misc_feature

15154..15720
/note="assembly_fragment"

misc_feature

15821..17135
/note="assembly_fragment"

misc_feature

17236..18343
/note="assembly_fragment"

misc_feature

18444..19570
/note="assembly_fragment"

Query Match

15.8%; Score 42.6; DB 2; Length 164490;

```

Best Local Similarity 48.5%; Pred. No. 0.71;
Matches 117; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 2 GGAGCATAAATCGTTATTAAAGATAATGGTATAGAACGAGCTTCGATGAATCAATGA 61
Db 32234 GTAGCAATTCATCAATGCTCAATGATACGATTCGCTTACTACAGAGAAAAG 32293
QY 62 TTTTATTGAGAAATCGTGGGACAGAGGGAAGAAAACAAAGCTCCCTGCGGAAG 121
Db 32294 CAATTCGTCTCTATGCAAAATTTATATAAGTAACCAAAAATGGCTACAGTGGGTGA 32353
QY 122 AATTCACAGGGTAAAGAGGCTGTTAAATGGCAATTCGCGCTTGGCAAAAT 181
Db 32354 GAGTCCTACAGAGATTTCTTCTAGCTCAGTTGGTACAGTATGCACTCAGAAATCAGG 32413
QY 182 TGAATTTCTACTATCCAGGGAACGAAAGGGTTACTTTTACTTTGGATTATCGAGAT 241
Db 32414 AGGATTATAGTTGTTCCAAAGAAAACAGATTTTGTGTCATGTTTGAATATAGATATAT 32473
QY 242 T 242
Db 32474 T 32474

RESULT 13
AC061968
LOCUS
DEFINITION
Homo sapiens chromosome 9 clone RP11-167K13 map 9, LOW-PASS
SEQUENCE SAMPLING.
AC061968
VERSION
HTG; HTGS PHASE0.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 63215)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 9, clone RP11-167K13
Unpublished
2 (bases 1 to 63215)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Girde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meltrim,J., Meneus,L., Mihova,T., Miranda,C., Miernga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu

```

```

----- Project Information
Center project name: L7248
Center clone name: 167_K_13

* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1
710: contig of 710 bp in length
810: gap of 100 bp
811: contig of 691 bp in length
1501: gap of 100 bp
1502: gap of 100 bp
1503: contig of 699 bp in length
2300: gap of 100 bp
2301: contig of 653 bp in length
3053: gap of 100 bp
3153: gap of 100 bp
3858: contig of 705 bp in length
3958: gap of 100 bp
4658: contig of 700 bp in length
4758: gap of 100 bp
5478: contig of 720 bp in length
5578: gap of 100 bp
6258: contig of 680 bp in length
6358: gap of 100 bp
7023: contig of 665 bp in length
7123: gap of 100 bp
7823: contig of 700 bp in length
7923: gap of 100 bp
8631: contig of 708 bp in length
8731: gap of 100 bp
9441: contig of 710 bp in length
9541: gap of 100 bp
10232: contig of 691 bp in length
10332: gap of 100 bp
11015: contig of 683 bp in length
11115: gap of 100 bp
11803: contig of 694 bp in length
11903: gap of 100 bp
12605: contig of 696 bp in length
12705: gap of 100 bp
13359: contig of 654 bp in length
13459: gap of 100 bp
14139: contig of 679 bp in length
14239: gap of 100 bp
14890: contig of 652 bp in length
14990: gap of 100 bp
15689: contig of 699 bp in length
15789: gap of 100 bp
16489: contig of 700 bp in length
16589: gap of 100 bp
17286: contig of 697 bp in length
17386: gap of 100 bp
18079: contig of 693 bp in length
18179: gap of 100 bp
18872: contig of 693 bp in length
18972: gap of 100 bp
19653: contig of 681 bp in length
19733: gap of 100 bp
20458: contig of 705 bp in length
20558: gap of 100 bp
21272: contig of 714 bp in length
21372: gap of 100 bp
22069: contig of 697 bp in length
22169: gap of 100 bp
22858: contig of 689 bp in length
22958: gap of 100 bp

```

sequencing vector: M13mp18; 55% of reads

Sequencing Vector: M13mp18; X02513; 99% of reads

Sequencing Vector:

----- Summary Statistics

Sequencing Vector:

Sequencing Vector: plasmid; plasmid_accession; 1% of reads
 Chemistry: Dye-primer; 2% of reads
 Assembly: Dye-terminator BigDye; 97% of reads
 Assembly program: Phrap; version 0.90319
 Consensus quality: 182658 bases at least Q40
 Consensus quality: 183349 bases at least Q30
 Consensus quality: 183837 bases at least Q20
 Insert size: 177966; agarose-fp
 Insert size: 185239; sum-of-contigs
 Quality coverage: 9.2x in Q20 bases; agarose-fp
 Quality coverage: 8.8x in Q20 bases; sum-of-contigs.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 3543: contig of 3543 bp in length
 * 3544 3643: gap of unknown length
 * 3644 39213: contig of 35570 bp in length
 * 39214 39313: gap of unknown length
 * 39314 87111: contig of 47798 bp in length
 * 87112 87211: gap of unknown length
 * 87212 185539: contig of 98328 bp in length.

FEATURES

source

1. 185539
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-140N7"
 /clone_lib="RPCI human BAC library 11"

misc_feature

1. 3543
 /note="assembly_name:Contig12
 clone_end:77"

misc_feature

3644. 39213
 /note="assembly_name:Contig13
 clone_end:SP6"

misc_feature

39314. 87111
 /note="assembly_name:Contig14"

misc_feature

87212. 185539
 /note="assembly_name:Contig15"

BASE COUNT 55079 a 38207 c 39983 g 52966 t 304 others

ORIGIN

Query Match 15.2%; Score 41; DB 2; Length 185539;
 Best Local Similarity 48.1%; Pred. No. 1.9;
 Matches 116; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 2 GGAGCATAAATCGTTATTAAAGATATGGTATAGGAACGAGCTTCGATGAATCAATGA 61
 Db 73259 GTAGCAITTGATCATTCCTCAATGATACGATTCGATGGCTTACTACAGAGGAAAG 73200
 QY 62 TTTTATTGAGATTCGGTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
 Db 73199 CAATTGGTCTCTATGCAATTTATTAAGTAATTAAGTAATTAAGTAATTAAGTAAT 73140
 QY 122 AATTCACGGGTAAAAAGCCCTTGGTAAATTCGCATTATTCGGCTTGGCAACAAAT 181
 Db 73139 GAGTCCTACAGGAAGATTCCTCTAGCTCAGTTCGATGATGACTCAGAAATCAGG 73080
 QY 182 TGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTTGGATTATGCAGAGAT 241
 Db 73079 AGGGATTTAGTGTTCCTCAAGGAAACAGATTTTGTCAATGATATAGATATAT 73020
 QY 242 T 242
 Db 73019 T 73019

RESULT 15

AC010279

LOCUS

DEFINITION

AC010279

AC010279

VERSION

AC010279.5

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Unpublished

JOURNAL

REFERENCE

AUTHORS

TITLE

DOE Joint Genome Institute

JOURNAL

REFERENCE

AUTHORS

TITLE

DOE Joint Genome Institute

JOURNAL

COMMENT

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www.shgc.stanford.edu

Quality: Phrap Quality >40 99.9% of Sequence;

Estimated Total Number of Errors is 0.1.

FEATURES

source

1. 125746
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTC-533D18"

BASE COUNT 36432 a 25088 c 26461 g 37765 t

ORIGIN

Query Match 14.6%; Score 39.4; DB 9; Length 125746;
 Best Local Similarity 47.7%; Pred. No. 5;
 Matches 115; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 2 GGAGCATAAATCGTTATTAAAGATATGGTATAGGAACGAGCTTCGATGAATCAATGA 61
 Db 96193 GTAGCAITTGATCATTCCTCAATGATACGATTCGATGGCTTACTACAGAGGAAAG 96252
 QY 62 TTTTATTGAGATTCGGTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
 Db 96253 CAATTGGTCTCTATGCAATTTATTAAGTAATTAAGTAATTAAGTAATTAAGTAAT 96312
 QY 122 AATTCACGGGTAAAAAGCCCTTGGTAAATTCGCATTATTCGGCTTGGCAACAAAT 181
 Db 96313 GAGTCCTACAGGAAGATTCCTCTAGCTCAGTTCGATGATGACTCAGAAATCAGG 96372
 QY 182 TGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTTGGATTATGCAGAGAT 241
 Db 96373 AGGGATTTAGTGTTCCTCAAGGAAACAGATTTTGTCAATGATATAGATATAT 96432
 QY 242 T 242
 Db 96433 T 96433

Search completed: November 15, 2003, 03:33:52
 Job time : 1194.78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 173.613 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-77

Perfect score: 269

Sequence: 1 CGGAGCATAAATCGTTATT.....GCAAGGTATTATCAACG 269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	269	100.0	269	AAV03585	Neisseria meningitidis
2	269	100.0	269	AAAI5389	Genomic fragment o
3	264.2	98.2	858	AAAI5296	DNA encoding a pol
4	246.6	91.7	855	ABZ37755	N. gonorrhoeae nuc
5	246.6	91.7	858	AAAI5322	DNA encoding a pol
6	246.6	91.7	891	ABZ11353	N. gonorrhoeae nuc
7	121.8	45.3	270	AAV03596	Neisseria meningitidis
8	121.8	45.3	270	AAAI5393	Genomic fragment o

C 9	41	15.2	13443	22	ABA19170	Human nervous syst
C 10	41	15.2	13443	22	AAK81300	Human immune/haema
C 11	39.4	14.6	13451	22	ABA19169	Human nervous syst
C 12	39.4	14.6	13451	22	AAK81299	Human immune/haema
C 13	36.8	13.7	4951	23	ABL13122	Drosophila melanog
C 14	36.8	13.7	5395	23	ABL15458	Drosophila melanog
C 15	36	13.4	2000	24	ABZ16977	Arabidopsis thalia
C 16	36	13.4	185695	25	ABQ77405	Human THBS1 DNA
C 17	35.4	13.2	294528	24	ABA03041	Listeria monocytog
C 18	34.8	12.9	1855	24	ABQ70544	Listeria monocytog
C 19	34.8	12.9	1855	24	ABQ71183	Listeria monocytog
C 20	34	12.6	381	22	AAF65964	Novel human polynu
C 21	33.8	12.6	733	22	AAK62701	Human immune/haema
C 22	33.6	12.5	7160	20	AAK86191	Nucleotide sequenc
C 23	33	12.3	1142	24	ABN89001	Human prostate spe
C 24	33	12.3	2729	22	AAK81142	Human immune/haema
C 25	32.8	12.2	29871	24	ABN86359	L. lactis multidru
C 26	32.6	12.1	1881	24	ABQ70528	Listeria monocytog
C 27	32.4	12.0	1664976	19	AAV21209	Methanococcus jann
C 28	32	11.9	1586	21	AAK36395	Arabidopsis thalia
C 29	32	11.9	6243	20	AAZ09598	Clostridium specie
C 30	32	11.9	30179	24	ABK47964	Human alcohol dehy
C 31	31.8	11.8	435	24	ABN90660	Staphylococcus epi
C 32	31.8	11.8	1355	25	ABZ37481	M13-III nucleotide
C 33	31.8	11.8	4583	25	ABV75375	Arabidopsis TES1 m
C 34	31.8	11.8	4590	25	ABV75374	Arabidopsis wild-t
C 35	31.8	11.8	6118	24	ABL33031	Human immune syste
C 36	31.8	11.8	13138	22	AAK65690	Human immune/haema
C 37	31.6	11.7	5033	25	ABZ09999	Haematopoietic cel
C 38	31.6	11.7	7847	23	ABL17450	Drosophila melanog
C 39	31.6	11.7	8033	22	AAK46307	Tumour suppressor
C 40	31.6	11.7	8033	24	ABL70169	Chemically treated
C 41	31.6	11.7	8033	24	AAK61116	Human gene regulat
C 42	31.6	11.7	8033	24	ABK31202	Signal transductio
C 43	31.4	11.7	281	24	ABQ55863	Human ovarian anti
C 44	31.4	11.7	581	21	AAK59135	Human secreted pro
C 45	31.4	11.7	884	22	AAH33290	Human colon cancer

ALIGNMENTS

RESULT 1
AAV03585
ID AAV03585 standard; DNA; 269 BP.

AC AAV03585;
DT 22-OCT-1998 (first entry)
DE Neisseria meningitidis DNA sequence C40.

N. gonorrhoeae; N. lactamica; chromosome Z2491; region 1; region 2;
region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
meningitis; ss.

OS Neisseria meningitidis.

PN WO9802547-A2.

PD 22-JAN-1998.

PF 11-JUL-1997; 97WO-FR01295.

PR 12-JUL-1996; 96FR-0008768.

(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
(SMIK) SMITHKLINE BEECHAM.

Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
Vinals C;

DR WPI; 1998-110594/10..

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX

PS Example 4; Page 118; 150pp; French.

XX

CC AA03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, rPA or C,

CC opC, porA, rotamase, sequence IC1106, IGA protease, pilin, pilC,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC *lactamica*, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX

SQ Sequence 269 BP; 95 A; 41 C; 65 G; 68 T; 0 other;

Query Match 100.0%; Score 269; DB 19; Length 269;

Best Local Similarity 100.0%; Pred. No. 5.5e-71;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATCAATCAATG 60

DB 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATCAATCAATG 60

QY 61 ATTTTATTTCGAGATCGGTCGGAACGAGAGGGAAGAAAACAAAGCTCCCGTCGGGAA 120

DB 61 ATTTTATTTCGAGATCGGTCGGAACGAGAGGGAAGAAAACAAAGCTCCCGTCGGGAA 120

QY 121 GAATTCGAACGGGTAAAGAGCGCTTCGTAATGCGATTTATTCGGGCTTGGCAACAAA 180

DB 121 GAATTCGAACGGGTAAAGAGCGCTTCGTAATGCGATTTATTCGGGCTTGGCAACAAA 180

QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240

DB 181 TTGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240

QY 241 TTCGAAGAAGCAAGGGTATTATCAACCG 269

DB 241 TTCGAAGAAGCAAGGGTATTATCAACCG 269

RESULT 2

AAA15389

ID AAA15389 standard; DNA; 269 BP.

XX AAA15389;

AC

XX

DT 04-SEP-2000 (first entry)

XX

DE Genomic fragment of *Neisseria meningitidis* Z2491.

XX

KW Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

XX

OS *Neisseria meningitidis*.

XX

PN WO200026375-A2.

XX

PD 11-MAY-2000.

XX

PF 28-OCT-1999; 99WO-FR02643.

XX

PR 30-OCT-1998; 98FR-0013693.

XX

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX

XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

PI Tinsley C, Perrin A;

XX

DR WPI; 2000-365622/31.

XX

PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic

PT or preventative vaccines and for diagnosis

XX

XX Claim 1; Page 40; 187pp; French.

XX

CC The present sequence represents a genomic fragment of *Neisseria*

CC meningitidis. The specification describes proteins that are specific

CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,

CC or their antigenic fragments, are used in vaccines to treat or protect

CC against *Neisseria* infections, particularly by *N. meningitidis*. The

CC polynucleotide sequences are also used for recombinant production of

CC the polypeptide and to produce attenuated *Neisseria* strains that

CC overexpress it, or express it in a non-toxic mutant form.

CC note: the present sequence is included in Claim 1, but it is those

CC sequences that do not include the present sequence that are actually

CC claimed.

XX

SQ Sequence 269 BP; 95 A; 41 C; 65 G; 68 T; 0 other;

Query Match 100.0%; Score 269; DB 21; Length 269;

Best Local Similarity 100.0%; Pred. No. 5.5e-71;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATCAATCAATG 60

DB 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATCAATCAATG 60

QY 61 ATTTTATTTCGAGATCGGTCGGAACGAGAGGGAAGAAAACAAAGCTCCCGTCGGGAA 120

DB 61 ATTTTATTTCGAGATCGGTCGGAACGAGAGGGAAGAAAACAAAGCTCCCGTCGGGAA 120

QY 121 GAATTCGAACGGGTAAAGAGCGCTTCGTAATGCGATTTATTCGGGCTTGGCAACAAA 180

DB 121 GAATTCGAACGGGTAAAGAGCGCTTCGTAATGCGATTTATTCGGGCTTGGCAACAAA 180

QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240

DB 181 TTGAATTTCTACTATCCAGGAAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240

QY 241 TTCGAAGAAGCAAGGGTATTATCAACCG 269

DB 241 TTCGAAGAAGCAAGGGTATTATCAACCG 269

RESULT 3

AAA15296

ID AAA15296 standard; DNA; 858 BP.

XX AAA15296;

AC

XX

DT 04-SEP-2000 (first entry)

XX

DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.

XX

KW Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

XX

OS *Neisseria meningitidis*.

XX

PN Key Location/Qualifiers

XX CDS 1..858

XX FT /*tag= a

XX

PN WO200026375-A2.

XX

PD 11-MAY-2000.

XX

```

XX 28-OCT-1999; 99WO-FR02643.
XX 30-OCT-1999; 98FR-0013693.
XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX (INMR ) INSERM INST NAT SANTE & RECH MEDICALE.
XX Autume L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
XX Tinsley C, Perrin A;
XX WPI; 2000-365622/31.
XX P-PSDB; AAY93265.
XX New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX or preventative vaccines and for diagnosis -
XX Claim 2; Page 51-52; 187pp; French.
XX The present sequence encodes a protein that is specific for pathogenic
XX strains of Neisseria. The polynucleotides, polypeptides, or their
XX antigenic fragments, are used in vaccines to treat or protect against
XX Neisseria infections, particularly by N. meningitidis. The
XX polynucleotide sequence is also used for recombinant production of
XX the polypeptide and to produce attenuated Neisseria strains that
XX overexpress it, or express it in a non-toxic mutant form.
XX Sequence 858 BP; 297 A; 145 C; 182 G; 234 T; 0 other;
SQ
Query Match 98.2%; Score 264.2; DB 21; Length 858;
Best Local Similarity 98.9%; Pred. No. 2.1e-69;
Matches 266; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGAGCATAAATCGTTATTAAAGATAATGTTAGGAACGAGCTTCGATGAATCAATG 60
Db 158 CGGAGCATAAATCGTTATTCAAGATAATGTTAGGAATGAGCTTCGATGAATCAATG 217
QY 61 ATTTTATTTCGAGATCGTTCGGAACAGAGCGGGAAGAAACAAAGCTCCCGTGGGAA 120
Db 218 ATTTTATTTCGAGATCGTTCGGAACAGAGCGGGAAGAAACAAAGCTTCCGTGGGAA 277
QY 121 GAATTCGAACGGGTAAAGAGCGCTTGTAAATTCGCATTATTTCGGCTTGGCAACAAA 180
Db 278 GAATTCGAACGGGTAAAGAGCGCTTGTAAATTCGCATTATTTCGGCTTGGCAACAAA 337
QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 240
Db 338 TTGAATTTCTACTATCCAGGAAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 397
QY 241 TTCGAAGACGACGGGTATTATCAACCG 269
Db 398 TTCGAAGACGACGGGTATTATCAACCG 426
RESULT 4
ABZ37755
ID ABZ37755 standard; DNA; 855 BP.
XX AC ABZ37755;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 99.
XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WO200279243-A2.
XX FD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB02069.

```

```

XX 12-FEB-2001; 2001GB-0003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX P-PSDB; ABP76785.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection -
XX Claim 6; Page 188; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX molecules of the invention.
XX Sequence 855 BP; 300 A; 151 C; 178 G; 226 T; 0 other;
SQ
Query Match 91.7%; Score 246.6; DB 25; Length 855;
Best Local Similarity 94.8%; Pred. No. 4.1e-64;
Matches 255; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 CGGAGCATAAATCGTTATTAAAGATAATGTTAGGAACGAGCTTCGATGAATCAATG 60
Db 158 CGGAGCATAAATCGTTATTAAAGATAATGTTAGGAATGAGCTTCGATGAATCAATG 217
QY 61 ATTTTATTTCGAGATCGTTCGGAACAGAGCGGGAAGAAACAAAGCTCCCGTGGGAA 120
Db 218 ATTTTATTTCGAGATCGTTCGGAACAGAGCGGGAAGAAACAAAGCTCCCGTGGGAA 277
QY 121 GAATTCGAACGGGTAAAGAGCGCTTGTAAATTCGCATTATTTCGGCTTGGCAACAAA 180
Db 278 GAATTCGAACGGGTAAAGAGCGCTTGTAAATTCGCATTATTTCAGGCTTGGCAACAAA 337
QY 181 TTGAATTTCTACTATCCAGGAAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 240
Db 338 TCGAATTCCTACTATCCAGGAAACGAAAGGGTTACTTTTACTTTGGATTATGCAGAGA 397
QY 241 TTCGAAGACGACGGGTATTATCAACCG 269
Db 398 TTAAGAAAAGTGAAGGTATTATCAACCG 426
RESULT 5
AAA15322
ID AAA15322 standard; DNA; 858 BP.
XX AC AAA15322;
XX DT 04-SEP-2000 (first entry)
XX DE DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX OS Neisseria gonorrhoeae.
XX FH Key Location/Qualifiers
XX CDS 1..858
XX FT /*tag= a
XX PN WO200026375-A2.
XX PD 11-MAY-2000.
XX

```

PF 28-OCT-1999; 99WO-FR02643.
 PR 30-OCT-1998; 98FR-0013693.
 XX (INRM) PASTEUR MERIEUX SERUMS & VACCINS SA.
 FA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Aujaume L, Bouchardon A, Renauld-Mongenle G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 XX WPI; 2000-365622/31.
 DR P-PSDB; AAY93292.
 XX
 PR New polypeptide specific for pathogenic Neisseria useful in therapeutic
 PT or preventative vaccines and for diagnosis -
 XX Claim 4; Page 134-135; 187pp; French.
 XX The present sequence encodes a protein that is specific for pathogenic
 CC strains of Neisseria. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC Neisseria infections, particularly by N. meningitidis. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated Neisseria strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 XX
 SQ Sequence 858 BP; 302 A; 151 C; 178 G; 227 T; 0 other;
 Query Match 91.7%; Score 246.6; DB 21; Length 858;
 Best Local Similarity 94.8%; Pred. No. 4.1e-64;
 Matches 255; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATGAAATCAATG 60
 DB 158 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAATGAGCTTCGATGAAATCAATG 217
 QY 61 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAACAAAGCTCCCGTGGGAA 120
 DB 218 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAACAAAGCTCCCGTGGGAA 277
 QY 121 GAATTCACACGGTAAAGAGGCTTCGTTAAATGCGATTATTCGGGCTTGGCAACAAA 180
 DB 278 GAATTCACACGGTAAAGAGGCTTCGTTAAATGCGATTATTCAGGCTTGGCAACAAA 337
 QY 181 TTGAATTTCTACTATCCAGGAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240
 DB 338 TCGAAATCTCTACTATCCAGGAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 397
 QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
 DB 398 TTAAAAAAGTGAGCGTATTATCAACCG 426
 RESULT 6
 ABZ41353
 ID ABZ41353 standard; DNA; 891 BP.
 AC ABZ41353;
 XX
 DT 07-MAR-2003 (first entry)
 DE N. gonorrhoeae nucleotide sequence SEQ ID 7295.
 XX Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX Neisseria gonorrhoeae.
 XX WO200279243-A2.
 XX 10-OCT-2002.
 XX 12-FEB-2002; 2002WO-IB02069.

PR 12-FEB-2001; 2001GB-0003424.
 XX (CHIR-) CHIRON SPA.
 PI Fontana MR, Pizza M, Masignani V, Monaci E;
 XX WPI; 2003-058415/05.
 DR P-PSDB; ABP80383.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection -
 XX Disclosure; Page 718; 815pp; English.
 XX The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
 CC molecules of the invention.
 XX
 SQ Sequence 891 BP; 316 A; 153 C; 186 G; 236 T; 0 other;
 Query Match 91.7%; Score 246.6; DB 25; Length 891;
 Best Local Similarity 94.8%; Pred. No. 4.1e-64;
 Matches 255; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAACGAGCTTCGATGAAATCAATG 60
 DB 194 CGGAGCATAAATCGTTATTAAAGATATGTTATAGGAATGAGCTTCGATGAAATCAATG 253
 QY 61 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAACAAAGCTCCCGTGGGAA 120
 DB 254 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAACAAAGCTCCCGTGGGAA 313
 QY 121 GAATTCACACGGTAAAGAGGCTTCGTTAAATGCGATTATTCGGGCTTGGCAACAAA 180
 DB 314 GAATTCACACGGTAAAGAGGCTTCGTTAAATGCGATTATTCAGGCTTGGCAACAAA 373
 QY 181 TTGAATTTCTACTATCCAGGAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 240
 DB 374 TCGAAATCTCTACTATCCAGGAACGAAAGGTTACTTTTACTTTGGATTATGCAGAGA 433
 QY 241 TTCGAAGAAGCAAGGTTATTATCAACCG 269
 DB 434 TTAAAAAAGTGAGCGTATTATCAACCG 462
 RESULT 7
 AAV03596/c
 ID AAV03596 standard; DNA; 270 BP.
 XX
 AC AAV03596;
 XX
 DT 22-OCT-1998 (first entry)
 XX
 DE Neisseria meningitidis DNA sequence E24.
 XX N. gonorrhoeae; N. lactamica; chromosome Z2491; region 1; region 2;
 KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
 XX meningitis; ss.
 XX Neisseria meningitidis.
 OS
 XX WO9802547-A2.
 XX
 PD 22-JAN-1998.
 XX
 PF 11-JUL-1997; 97WO-FR01295.
 XX
 PR 12-JUL-1996; 96FR-0008768.

```

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (SMIK ) SMITHKLINE BEECHAM.
XX
XX Nassif X, Tinsley C, Achman M, Merker P, Ruelle J;
PI Vinals C;
XX
XX WPI; 1998-110594/10. .
XX
XX Genes present in Neisseria meningitidis but not other Neisseria
PT species - and related host cells' RNA, anti-sense sequences.
PT polypeptide(s) and antibodies, useful for diagnosing Neisseria
PT meningitidis infection and in protective vaccines
XX
XX Example 4; Page 124; 150pp; French.
XX
XX AA03575-606 represent sequences that are present in Neisseria
CC meningitidis and N. gonorrhoeae but not in N. lactamica, except for the
CC genes involved in biosynthesis of the capsule polysaccharide, lfpA or C,
CC opc, porA, rotamase, sequence IC106, iga protease, pillin, pilC,
CC proteins which bind transferrin and opacity proteins. The DNA sequences
CC are responsible for the differences in pathogenicity between N.
CC meningitidis and N. gonorrhoeae, specifically they include the genes that
CC allow N. meningitidis to cross the blood-brain barrier. DNA sequences
CC common to N. meningitidis and N. gonorrhoeae, but absent from N.
CC lactamica, are responsible for colonisation and penetration of the
CC mucosa. The DNA sequences can be used to produce probes and primers, and
CC antibodies produced against the encoded proteins are used in standard
CC hybridisation/immunoassay processes for diagnosis of N. meningitidis
CC infection, particularly meningitis.
XX
XX Sequence 270 BP; 72 A; 67 C; 42 G; 89 T; 0 other;
XX
XX Query Match 45.3%; Score 121.8; DB 19; Length 270;
XX Best Local Similarity 98.4%; Pred. No. 8.2e-27;
XX Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CGGAGCATAAAATCGTTATTAAAGATATATGTTATAGGACGAGCTTCGATGAAATCAATG 60
XX Db 125 CGGAGCATAAAATCGTTATTAAAGATATGTTATAGGAAATCAATG 66
XX
XX QY 61 ATTTTATTTGAGAAATCGTTCGGAACAGAGGGAAGAAAACAGCCTCCCGTGCAGAA 120
XX Db 65 ATTTTATTTGAGAAATCGTTCGGAACAGAGGGAAGAAAACAGCCTCCCGTGCAGAA 6
XX
XX QY 121 GAATT 125
XX Db 5 GAATT 1
XX
XX RESULT 8
XX AAAL5393/C
XX ID AAAL5393 standard; DNA; 270 BP.
XX
XX AC AAAL5393;
XX
XX XX 04-SEP-2000 (first entry)
XX
XX Genomic fragment of Neisseria meningitidis Z2491.
XX
XX Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
XX Neisseria meningitidis.
XX
XX WO200026375-A2.
XX
XX 11-MAY-2000.
XX
XX 28-OCT-1999; 99WO-FR02643.
XX
XX 30-OCT-1998; 98PR-0013693.
XX

```

PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0255719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02559678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PI useful for preventing, diagnosing and/or treating nervous system
 XX cancers and metastases -

PS Disclosure; SEQ ID NO 11501; 1701bp + Sequence Listing; English.

The invention relates to novel genes (AB211004-AB211534) and proteins (AB214678-AB218001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 13443 BP; 4183 A; 2486 C; 2503 G; 4271 T; 0 other;

Query Match 15.2%; Score 41; DB 22; Length 13443;
Best Local Similarity 48.1%; Pred. No. 0.045;
Matches 116; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 2 GGACATAAAATCGTTATTAAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATGA 61
Db 6508 GTAGCATTTGATCAATTCGACTTCATGATACCATTTGGATGGCCTTACTACAGAGGAAAG 6549

Qy 62 TTTTATTATGAGATTCGGTTCGACACAGAGGGAAGAAAACAACGCTCCCGTCGCGAAG 121
Db 6548 CAATTGGTGCTCTATGCAAAATTTATTAAGTAAACCAAGAAATGCCTACAGGTGGGTGA 6489

Qy 122 AATTCCAAACGGGTAAAAAAGCCCTTGGTAAATTGSCATTATTCGGGCTTCGCAACAAAT 181
Db 6488 GAGTCCTACAGGAAGATCTCTCTAGCTCAGTTGTCAGTATGTGACTCAGAAATCAGG 6429

Qy 182 TGAATTTCTACTATCCAGGGAACGAAAGGGTTACTTTTACTTTGATTATGCAGAGAT 241
Db 6428 AGGATTATTAGTGTTCCTCAAGAAAAACAGATTTTGTTCATGCTTTAGAAATAGATATAT 6369

Qy 242 T 242
Db 6368 T 6368

RESULT 10
AAK81300/c
ID AAK81300 standard; DNA; 13443 BP.
AC AAK81300;
XX
XX
XX 07-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36112.
DE
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX Cytostatic; Gene therapy; vaccines; metastasis; ds.
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0185974.
XX 17-MAR-2000; 2000US-0190076.

CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 13443 BP; 4183 A; 2486 C; 2503 G; 4271 T; 0 other;

 Query Match 15.2%; Score 41; DB 22; Length 13443;
 Best Local Similarity 48.1%; Fred. No. 0.045;
 Matches 116; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

 QY 2 GGAGCATAAATCGTTATTAAAGATAATGGTATAGGACGAGCTTCGATGAAATCAATGA 61
 DB 6608 GTAGCATTTGATCATTTGACATGATACGATTGGATGGCTTACTACAGAGGAAAG 6549

 QY 62 TTTTATTGAGATCGTTCGACAGAGGAGGAGAGAAACAGCTCCCGTCGGGAAG 121
 DB 6548 CAATTGGTGTCTATGCAAAATTTATTAAGTAACACGAGAAATGCGCTACAGTGGGTGA 6489

 QY 122 AATTCACAGGGTAAAGGCTTGTAAATTTGGCAATTTTCGGGCTTGCAACAAAT 181
 DB 6488 GAGTCTACAGAGATTTCTTCTAGCTCAGTGTGACAGTATGCTCAGAAATCAGG 6429

 QY 182 TGAATTTCTACTATCCAGGGAACGAAAGGTTACTTTTACTTTGATTTGACAGAT 241
 DB 6428 AGGGATTAGTGTTCCTCAAGAAAAACAGATTTTGTATGCTTTAGAAATATAGATAT 6369

 QY 242 T 242
 DB 6368 T 6368

 RESULT 11
 ABA19169/c
 ID ABA19169 standard; DNA; 13451 BP.
 XX
 AC ABA19169;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 11500.
 XX
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
 KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
 KW antitubercular; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR

20-OCT-2000; 2000US-0240960.
 20-OCT-2000; 2000US-0241221.
 20-OCT-2000; 2000US-0241785.
 20-OCT-2000; 2000US-0241786.
 20-OCT-2000; 2000US-0241787.
 20-OCT-2000; 2000US-0241808.
 20-OCT-2000; 2000US-0241809.
 20-OCT-2000; 2000US-0241826.
 20-OCT-2000; 2000US-0241827.
 01-NOV-2000; 2000US-0246474.
 08-NOV-2000; 2000US-0246475.
 08-NOV-2000; 2000US-0246476.
 08-NOV-2000; 2000US-0246477.
 08-NOV-2000; 2000US-0246478.
 08-NOV-2000; 2000US-0246523.
 08-NOV-2000; 2000US-0246524.
 08-NOV-2000; 2000US-0246525.
 08-NOV-2000; 2000US-0246526.
 08-NOV-2000; 2000US-0246527.
 08-NOV-2000; 2000US-0246528.
 08-NOV-2000; 2000US-0246532.
 08-NOV-2000; 2000US-0246609.
 08-NOV-2000; 2000US-0246610.
 08-NOV-2000; 2000US-0246611.
 08-NOV-2000; 2000US-0246613.
 17-NOV-2000; 2000US-0249207.
 17-NOV-2000; 2000US-0249208.
 17-NOV-2000; 2000US-0249209.
 17-NOV-2000; 2000US-0249210.
 17-NOV-2000; 2000US-0249211.
 17-NOV-2000; 2000US-0249212.
 17-NOV-2000; 2000US-0249213.
 17-NOV-2000; 2000US-0249214.
 17-NOV-2000; 2000US-0249215.
 17-NOV-2000; 2000US-0249216.
 17-NOV-2000; 2000US-0249217.
 17-NOV-2000; 2000US-0249218.
 17-NOV-2000; 2000US-0249219.
 17-NOV-2000; 2000US-0249224.
 17-NOV-2000; 2000US-0249245.
 17-NOV-2000; 2000US-0249264.
 17-NOV-2000; 2000US-0249265.
 17-NOV-2000; 2000US-0249297.
 17-NOV-2000; 2000US-0249299.
 17-NOV-2000; 2000US-0249300.
 01-DEC-2000; 2000US-0250160.
 01-DEC-2000; 2000US-0250391.
 05-DEC-2000; 2000US-0251030.
 05-DEC-2000; 2000US-0251989.
 05-DEC-2000; 2000US-0256719.
 06-DEC-2000; 2000US-0251479.
 08-DEC-2000; 2000US-0251856.
 08-DEC-2000; 2000US-0251868.
 08-DEC-2000; 2000US-0251869.
 08-DEC-2000; 2000US-0251989.
 11-DEC-2000; 2000US-0251990.
 08-DEC-2000; 2000US-0254097.
 05-JAN-2001; 2001US-0259676.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 Disclosure; SEQ ID NO 36112; 3071pp + Sequence Listing; English.
 AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 amino acid sequences given in AAK52170 to AAK91921. (I) have cytostatic
 activity, and can be used in gene therapy and vaccine production. (I)

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227709.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure; SEQ ID NO 11500; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 13451 BP; 4193 A; 2486 C; 2502 G; 4270 T; 0 other;

Query Match 14.6%; Score 39.4; DB 22; Length 13451;
 Best Local Similarity 47.7%; Pred. No. 0.14;
 Matches 115; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 2 GGACATAAATCCTTATTAAAGTATATGTTAGGAACGAGCTTCGATGAATCAATGA 61
 Db 6612 GTAGCATTTGATCAATTCACCTTCAATGATGATGATGCGCTTACTACAGAGGAAAG 6553
 QY 62 TTTTATTGAGATCGGTGGAAACAGAGGAGAAACACAGCTCCCGTCGGGAAG 121
 Db 6552 CAATGGTGCTTATGCAAAATTTATATAGTAAACAGAAATGCGCTACAGTGGGATGA 6493
 QY 122 AATTCACAGGGTAAAGAGCCCTTGTTAAATGTCATTTTCGGCTTGGCAACAAAT 181
 Db 6492 GAGTCCTACAGGAAGATTTCTTCTAGCTCAGTTGGTACAGTATGTGACTCAGAAATCAGG 6433
 QY 182 TGAATTTCTACTCCAGGGAACAGAAAGGTTTACTTTTGGATTATCGAGAT 241
 Db 6432 AGGATTTAGTGTTCACAGAAACACAGATTTTGTCATGCTTACGATATAGATATAT 6373
 QY 242 T 242
 Db 6372 T 6372

RESULT 12

AAK81299/c
 ID AAK81299 standard; DNA; 13451 BP.

AC AAK81299;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36111.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.

PT interactions -
 PS Claim 1; SEQ ID NO 33848; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 4951 BP; 1376 A; 1121 C; 1065 G; 1389 T; 0 other;
 Query Match 13.7%; Score 36.8; DB 23; Length 4951;
 Best Local Similarity 61.5%; Pred. No. 0.62;
 Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 QY 1 CGGAGCATAAATCGTTTATTAAAGATAATCGTTATAGGACGAGCTTCGATGAATCAATG 60
 DB 3962 CCGAGCATACATCGTTTATTAGTTATATGTTAAAGAGATAATTTGTTTAAATATATG 4021
 QY 61 ATTTTATTTCGAGATCGGTCGGAACAGAGGGAAG 96
 DB 4022 ATATATTTCGAGTATACGGTTACAGAGTGCATG 4057
 RESULT 14
 ABL15458
 ID ABL15458 standard; cDNA; 5395 BP.
 XX
 AC ABL15458;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 40856.
 XX
 KW *Drosophila*; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS *Drosophila melanogaster*.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB71355.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 40856; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 5395 BP; 1528 A; 1109 C; 1141 G; 1617 T; 0 other;
 Query Match 13.7%; Score 36.8; DB 23; Length 5395;
 Best Local Similarity 61.5%; Pred. No. 0.63;
 Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 QY 1 CGGAGCATAAATCGTTTATTAAAGATAATCGTTATAGGACGAGCTTCGATGAATCAATG 60
 DB 441 CCGAGCATACATCGTTTATTAGTTATATGTTAAAGAGATAATTTGTTTAAATATATG 500
 QY 61 ATTTTATTTCGAGATCGGTCGGAACAGAGGGAAG 96
 DB 501 ATATATTTCGAGTATACGGTTACAGAGTGCATG 536
 RESULT 15
 ABL216977/c
 ID ABL216977 standard; DNA; 2000 BP.
 XX
 AC ABL216977;
 XX
 DT 21-JAN-2003 (first entry)
 XX
 DE *Arabidopsis thaliana* stress regulated gene SEQ ID NO 4782.
 XX
 KW *Arabidopsis thaliana*; plant; gene; stress; transgenic; ds.
 XX
 OS *Arabidopsis thaliana*.
 XX
 PN WO200216655-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 24-AUG-2001; 2001WO-US26685.
 XX
 PR 24-AUG-2000; 2000US-227866P.
 PR 26-JAN-2001; 2001US-264647P.
 PR 22-JUN-2001; 2001US-300111P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper JF, Krepis J, Wang X, Zhu T;
 XX
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed
 PT and producing plants with increased tolerance to these abiotic stresses
 PT -
 XX
 PS Claim 14; SEQ ID NO 4782; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising:
 CC (a) contacting nucleic acid representative of expressed polynucleotides
 CC in the plant cell with an array or probes representative of the plant
 CC cell genome; and
 CC (b) detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an *Arabidopsis thaliana* stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.

Search completed: November 15, 2003, 00:35:27
Job time : 176.613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1794.52 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-77
Perfect score: 269
Sequence: 1 CGGAGCATAAATCGTTATT.....GCAAGGCTATTATCAACG 269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.6	15.1	1101	29	CNS016TL
2	39.2	14.6	1201	9	AL546670
3	38.8	14.4	289	29	T09518
4	38.6	14.3	797	13	EX437575

5	38.6	14.3	1201	9	AL550892
6	38.4	14.3	1201	13	BX358323
7	37.6	14.0	1201	13	BX364341
8	37.2	13.8	635	13	BO091774
9	37.2	13.8	859	29	CNS01227
10	37	13.8	1186	13	EX447889
11	36.8	13.7	521	10	AW942932
12	36.6	13.6	763	9	AJ449657
13	36.6	13.6	1201	9	AL534914
14	36.6	13.6	1310	29	CNS015ZW
15	36.4	13.5	895	28	AZ674925
16	36.2	13.5	1048	13	BX363929
17	36	13.4	512	28	BZ114936
18	36	13.4	1101	29	CNS00BK3
19	36	13.4	1173	29	CC280154
20	35.6	13.2	750	28	BH058154
21	35.2	13.1	505	28	AZ144133
22	35.2	13.1	698	28	AQ078372
23	35.2	13.1	723	13	BU006834
24	35.2	13.1	1101	29	CNS0039G
25	35.2	13.1	1101	29	CNS00DDL
26	35.2	13.1	1201	9	AL562196
27	35	13.0	458	28	AZ133779
28	34.8	12.9	451	14	CA410343
29	34.6	12.9	628	28	BZ121223
30	34.6	12.9	718	28	BZ023545
31	34.6	12.9	917	28	AZ680900
32	34.6	12.9	941	13	BX432938
33	34.6	12.9	983	13	BX356954
34	34.6	12.9	1048	29	CNS01412
35	34.6	12.9	1201	13	BX397003
36	34.4	12.8	441	9	AJ485891
37	34.4	12.8	527	9	AW616473
38	34.4	12.8	593	10	AW934374
39	34.4	12.8	800	13	BZ223445
40	34.4	12.8	847	13	BX453737
41	34.4	12.8	1201	13	BX355540
42	34.4	12.8	1201	29	CNS016FT
43	34.2	12.7	384	13	BY199244
44	34.2	12.7	404	13	BI341230
45	34.2	12.7	446	13	BU754325

ALIGNMENTS

RESULT 1
CNS016TL
LOCUS

DEFINITION

CNS016TL 1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN16121 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

AL107187

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

GI:5626043

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :

Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the European Drosophila Genome Project (EDGP) -

http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CEPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC

Project Grant. The DNA was prepared from embryos by Alain Bucheton

and Genevieve Payan. It has been constructed in the vector

AL550892 AL550892
BX358323 BX358323
BX364341 BX364341
BO091774 UMN21H10
AL101017 Drosophila
EX447889 EX447889
AJ942932 LD2090.3
AJ449657 AJ449657
AL534914 AL534914
AL106118 Drosophila
AZ674925 ENTJ21TR
BX363929 BX363929
BZ114936 CH230-255
AL069292 Drosophila
CC280154 CH261-24H
BH058154 RPT-24-3
AZ144133 SP 0019 A
AQ078372 CIT-HSP2
BU006834 QGH12113
AL063921 Drosophila
AL069883 Drosophila
AL562196 AL562196
AZ133779 OSUNB010
CA410343 614 F -P
BZ121223 CH270-411
BZ023545 ced89d05
AZ680900 ENTGU83TR
BX432938 BX432938
BX356954 BX356954
AL103568 Drosophila
BX397003 BX397003
AI485891 EST244212
AW616473 EST322884
AW934374 EST360217
BU223445 603948433
BX453737 BX453737
BX355540 BX355540
AL106691 Drosophila
BY199244 BY199244
BI341230 368603 MA
BU754325 UI-1-BB1P

```

FEATURES
  source
    pBelobAC11.
      Location/Qualifiers
        1. 1101
          /organism="Drosophila melanogaster"
          /mol_type="genomic DNA"
          /db_xref="taxon:7227"
          /clone="BACN16L21"
          /clone_lib="DrosBAC"
          /plasmid="pBelobAC11"
          /note="end : T7"
BASE COUNT      402 a 153 c 67 g 213 t 266 others
ORIGIN
Query Match      15.1%; Score 40.6; DB 29; Length 1101;
Best Local Similarity 32.5%; Pred. No. 2.3;
Matches 79; Conservative 64; Mismatches 99; Indels 1; Gaps 1;
QY 16 TTATTAAGATAAATGGTATAGGACGAGCTTCGATGAATCAATGATTTTATTGAGAA 75
Db 823 TTTTSTAWATAAAMCASAAMAAASVASSHTCSAATTAANAANKTWTWTTAAATATKGA 882
QY 76 TCGGTCGGAACAGAGGAGAAACAAACAGCCTCCCGTCGCGGAAGAAATCCAAACGGTA 135
Db 883 TAGCGGGGAAATAACGAKASSDRAACGSSBSGGGAATKASAAVATTATTSAWCSTAA 942
QY 136 AAAAAGCCCTTGGTAATGTCATATTCGGGCTTGGCAAC-AAATTTGAATTTCTACT 194
Db 943 AADADSKSSGTFBKWSTGGKRTTATTADACBMAAACCCCBTAACBKAACOMMWA 1002
QY 195 ATCCAGGGAACGAAGAGGTTACTTTTACTTTGATTATGAGAGATTCCGAAGAACAAG 254
Db 1003 TTTWAAAAAAMAMACCVTTAASTWAAAVTTTBTATSSGVBSTTSSMSDSTSSSS 1062
QY 255 GGT 257
Db 1063 SST 1065

RESULT 2
AL546670      1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
  AL546670 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
  clone CS0D1029YD16 5-PRIME, mRNA sequence.
ACCESSION
  AL546670
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1201)
  Li W.B., Gruber C., Jesses, J. and Polayes D.
  Full-length cDNA libraries and normalization
  Unpublished
  On Feb 15, 2001 this sequence version replaced gi:12880010.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1029YD08QF1.
FEATURES
  source
    1. 1201
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0D1029YD16"
      /tissue_type="PLACENTA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-cligo(dt)
      primer. Five prime end enriched, double-strand cDNA was

```

```

digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      302 a 255 c 226 g 296 t 122 others
ORIGIN
Query Match      14.6%; Score 39.2; DB 9; Length 1201;
Best Local Similarity 36.4%; Pred. No. 5.7; Mismatches 37; Indels 0; Gaps 0;
Matches 68; Conservative 37; Mismatches 37; Indels 0; Gaps 0;
QY 63 TTTTATTTGAGATCGTCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
Db 915 WTTTTTTTTTAAAAAGGKYCYVTWAAWRKTTTWWAAAAAAGVMDGGSSGKTTTAAA 974
QY 123 ATTCACGGGTAAAAAAGCGCTTGGTAATTCGATTAATTCGGCTTGGCAACAAAT 182
Db 975 AAKGARKARGAAAWANWAKTTTGGKTTTTTTAAARWCTHTTTTGGGGRAAAATG 1034
QY 183 GAAATTTCTACTATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 242
Db 1035 RAAAAAMCWAARGITTHMSBYAAGGAWAWNTTTTDTKTTTTTAAAAAASAAHBS 1094
QY 243 CGAAGAA 249
Db 1095 GSKGRAA 1101

T09518      289 bp DNA linear GSS 05-DEC-2000
LOCUS
DEFINITION
  T09518 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
  clone 0113m, genomic survey sequence.
ACCESSION
  T09518
VERSION
  T09518.1 GI:319350
KEYWORDS
  GSS.
SOURCE
  Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM
  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
  1 (bases 1 to 289)
  Reddy, G.R., Chakrabarti, D., Schuster, S.M., Ferl, R.J., Almira, E.C.
  and Dame, J.B.
  Gene sequence tags from Plasmodium falciparum genomic DNA fragments
  prepared by the genome activity of mung bean nuclease
  Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)
JOURNAL
  MEDLINE
  94052193
COMMENT
  Other GSSs: 0113m3
  Contact: Dame JB
  Dept. of Pathobiology, College of Veterinary Medicine
  University of Florida
  2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
  Tel: 352 392 4700
  Fax: 352 392 9704
  Email: damej@mail.vetmed.ufl.edu
  Seq primer: T7
  Class: Shotgun.
FEATURES
  Location/Qualifiers
    1. 289
      /organism="Plasmodium falciparum"
      /mol_type="genomic DNA"
      /db_xref="taxon:5833"
      /clone="0113m"
      /lab_host="E. coli XL1-Blue"
      /clone_lib="gmbpFHB3.1, G. Roman Reddy"
      /note="Vector: pBluescript SK(+); Genomic DNA, from
      asynchronous blood stage parasites of the cloned Honduran
      HB3 isolate cultured in vitro, was digested with mung bean
      nuclease in the presence of 30% formamide at 50°C (Vernick
      K.D., Imberski, R.B., and McCutchan, T.F. 1989. Nucleic
      Acids Research 16:6831-6896). The ends of the fragments
      were polished using T4 DNA polymerase, and the fragments
      were ligated to EcoR V-cleaved and dephosphorylated
      pBluescript SK(+). Recombinant plasmids transformed E.
      coli XL1-Blue."

```


BASE COUNT	117 a	29 c	31 g	100 t	12 others
ORIGIN					
Query Match	14.4%;	Score 38.8;	DB 29;	Length 289;	
Best Local Similarity	48.4%;	Pred. No. 5.8;			
Matches	91;	Conservative	0;	Mismatches	97;
Indels					
Gaps					
QY	4	AGCATAAATCGTTTAAAGATAATGTTATAGGAACAGAGCTTCGATGAATCAATGATT	63		
DB	278	AGCANTGANNATATATATAAGAAATANTATTGGTGAAATATTTTTGATATGATTAATC	219		
QY	64	TTTATTGAGATCGGTCCGACAGAGGAGAGAAACCAAGCTCCCGTCGGCGAAGAA	123		
DB	218	NTTTTTCAAATCATTATGTAAATATGTATGTTTATTAACTTACCTTCAGATGCA	159		
QY	124	TTCCACGGGTAAAAAGCCCTTGTTAAATGGCATTATTCGGCGTTGGCAACAAAATTG	183		
DB	158	CTAANAATTTAAAAAGTCTTCATATTACNVACATATATTTTNGTGTCCAAATCCTTT	99		
QY	184	AAATTTCT	191		
DB	98	AAATTTT	91		
RESULT 4					
LOCUS	BX437575/c	797 bp	mRNA	linear	EST 15-MAY-2003
DEFINITION	BX437575 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP007YF09				
ACCESSION	BX437575				
VERSION	BX437575.1 GI:30771584				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
COMMENT	Contact: Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10594.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSOCAP007CC05NP1&cluster=10594.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOCAP007CC05NP1. Location/Qualifiers 1. .797 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOCAP007YF09" /tissue_type="THYMUS" /clone_lib="Homo sapiens THYMUS" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."				
BASE COUNT	101 a	142 c	102 g	288 t	164 others
ORIGIN					
Query Match	14.3%;	Score 38.6;	DB 13;	Length 797;	
Best Local Similarity	26.7%;	Pred. No. 7.7;			
Matches	65;	Conservative	75;	Mismatches	103;
Indels					
Gaps					
Y	7	ATAAATCGTTTAAAGATAATGTTATAGGAACAGAGCTTCGATCAATCAATGATTTT	66		

```

Db      213 TATAGAAAAA... 1201 bp mRNA linear EST 02-MAY-2003
LOCUS
DEFINITION BX358323 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1036YM12 3-PRIME, mRNA sequence.
ACCESSION BX358323
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1036YG06NP1.
FEATURES
source
1..1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036YM12"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 356 a 225 c 151 g 115 t 354 others
ORIGIN
Query Match 14.3%; Score 38.4; DB 13; Length 1201;
Best Local Similarity 21.3%; Pred. No. 9.3;
Matches 32; Conservative 65; Mismatches 53; Indels 0; Gaps 0;

QY 9 AAAATCGTTAATAGTAATGGTATAGGAACGAGCTTCGATGAATCAATGATTTTAT 68
Db 1008 ATATKATATATKAAATAATKMAHAATKMKMMKMKKNDKXMMMMMAAKMMKYYDD 1067
QY 69 TTGAGATCGTCGGAAGGAGAGAAAAAACAAGCTCCCGTCGGGGAAGAAATCCA 128
Db 1068 KKNWGNAMCMKGMNMGAAKRAAMWMAKMKDKRAAKARGKKKKKKADKKX 1127
QY 129 ACGGGTAAAAAGCGCTTGGTAAATTCGCA 158
Db 1128 AKKKRMAAAAKKAKKKKKADKKKAA 1157

RESULT 7

```

```

BX364341 1201 bp mRNA linear EST 05-MAY-2003
LOCUS
DEFINITION BX364341 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL012Y017 5-PRIME, mRNA sequence.
ACCESSION BX364341
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3185.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL012AH09QP1&cluster=3185.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL012AH09QP1.
FEATURES
source
1..1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL012Y017"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="RAMOS CELL LINE"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 283 a 297 c 308 g 217 t 96 others
ORIGIN
Query Match 14.0%; Score 37.6; DB 13; Length 1201;
Best Local Similarity 39.4%; Pred. No. 15;
Matches 86; Conservative 37; Mismatches 93; Indels 2; Gaps 1;

QY 19 TTAAGATAATGGTATAGGAACGAGCTTCGATGAATCAATGATTTTATTGAGATCG 78
Db 835 BYAAAAAAMWMAAAAAAAGGKCMAGATTAASAAHAWTTTTTTTAAAAATGG 894
QY 79 GTCCGACAGAGGGAAGAAAAACAGGCTCCCGTCGGGAGGATTCACACGGGTAAAA 138
Db 895 TCC-AAAAAAAKVGAATAAACHAAGAGVGSCTCTTTTAAAGVSTAYGGGAAAM 952
QY 139 AAGCGCTTGGTAAATTCGATTTTCGGGCTTGGCAACAAAATTCGAAATTTCTACTCC 198
Db 953 TSYACTTGGTTTTTGAGGACATTTTGTGTTTAAATKGRMAAYACVMAAATTA 1012
QY 199 AGGGAACGAAAGGTTACTTTTACTTTGGATTATGCA 236
Db 1013 AVCYAGTAATAAAWTTTADTGTAAATTTTAAAYAVTGMA 1050

RESULT 8
LOCUS
DEFINITION BQ091774 Canine Brain cDNA Library Canis familiaris CDNA 5' similar
to Canis familiaris mitochondrion, complete genome, mRNA sequence.
ACCESSION BQ091774
VERSION EST.
KEYWORDS

```

```

SOURCE
ORGANISM      Canis familiaris (dog)
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
AUTHORS       Roberts, M.C., Hendrickson, J.A., Hoffmann, D.E., Flickinger, G.H.,
               Rutherford, M.S. and Mickelson, J.R.
TITLE         University of Minnesota Canine Brain EST Project
JOURNAL       Unpublished
COMMENT       Contact: Mickelson, J.R.
               Veterinary Pathobiology
               University of Minnesota
               1988 Fitch Avenue, University of Minnesota, St. Paul, MN 55108, USA
               Tel: 612 624 1246
               Fax: 612 625 0204
               Email: mickel001@umn.edu
               Seq primer: M13 Reverse
               Location/Qualifiers
               1..635
               /organism="Canis familiaris"
               /mol_type="mRNA"
               /db_xref="taxon:9615"
               /sex="Male"
               /clone_lib="Canine Brain cDNA Library"
               /note="Organ: Brain; Vector: pSPORT1 (Gibco BRL); Site: 1:
               Not1; Site 2: SalI; Tissue was taken from the frontal,
               occipital, temporal and parietal lobes, olfactory bulb,
               hippocampus, cerebellum, thalamus, hypothalamus, midbrain
               pons, and medulla."
BASE COUNT    170 a 122 c 128 g 207 t 8 others
ORIGIN
               Query Match      13.8%; Score 37.2; DB 13; Length 635;
               Best Local Similarity 55.2%; Pred. No. 18;
               Matches 69; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 123 ATCCACGGGTAAAGAGCGCTTGGTAAATGCGATATTCGGGCTTGCACAAATTT 182
Db 123 ATCCACGGGTAAAGAGCGCTTGGTAAATGCGATATTCGGGCTTGCACAAATTT 182

QY 380 ATCCACGGGAGTAAAGATTTAGTTGCTGCAACACTTCATGGAGGCAATATTTAA 439
Db 380 ATCCACGGGAGTAAAGATTTAGTTGCTGCAACACTTCATGGAGGCAATATTTAA 439

QY 183 GAAATTTCTACTCCAGGGAACAAAGGTTTACTTTTACTTTGGATTATGCAGAGATT 242
Db 183 GAAATTTCTACTCCAGGGAACAAAGGTTTACTTTTACTTTGGATTATGCAGAGATT 242

QY 440 TGATCTTCANCTNGCTATGAGCTTTAAGGGTTATTTTTTACCAAGTGGNGGTTA 499
Db 440 TGATCTTCANCTNGCTATGAGCTTTAAGGGTTATTTTTTACCAAGTGGNGGTTA 499

QY 243 CGAAG 247
Db 243 CGAAG 504

RESULT 9
CNS01227/c    859 bp    DNA    linear    GSS 26-JUL-1999
LOCUS         Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION    BACN07C22 of DrosBAC library from Drosophila melanogaster (fruit
               fly), genomic survey sequence.
ACCESSION     AL101017
VERSION       AL101017.1 GI:5612628
KEYWORDS      GSS.
SOURCE        Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
               1 (bases 1 to 859)
               Genoscope.
REFERENCE
AUTHORS       Direct Submission
TITLE         Submitted (23-JUL-1999); Genoscope - Centre National de Sequencage :
               BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
               - Web : www.genoscope.cns.fr)
JOURNAL       Determination of this BAC-end sequence was carried out as part of a
               collaboration with the European Drosophila Genome Project (EDGP) -
               http://www.edgp.sbi.ac.uk -. This Drosophila melanogaster BAC
               library (Dros BAC) was made by Alain Billaud at CEPH (Centre

```

```

d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source        1..859
               /organism="Drosophila melanogaster"
               /mol_type="genomic DNA"
               /db_xref="taxon:7227"
               /clone="BACN07C22"
               /clone_lib="DrosBAC"
               /plasmid="pBelOBAC11"
               /note="end: T7"
BASE COUNT    235 a 149 c 100 g 300 t 75 others
ORIGIN
               Query Match      13.8%; Score 37.2; DB 29; Length 859;
               Best Local Similarity 41.9%; Pred. No. 19;
               Matches 65; Conservative 22; Mismatches 68; Indels 0; Gaps 0;

QY 7 ATAAATCGTTATTAAAGATAATGTTATAGGACAGCTTCGATGAATCAATGATTTT 66
Db 7 ATAAATCGTTATTAAAGATAATGTTATAGGACAGCTTCGATGAATCAATGATTTT 66

QY 67 ATTGAGAATCGTTCGGAACAGAGGAAGGAAACAAAGCCCTCCCGTCGCGAAGATTC 126
Db 67 ATTGAGAATCGTTCGGAACAGAGGAAGGAAACAAAGCCCTCCCGTCGCGAAGATTC 126

QY 750 ATATAAATGTAKTGTAGAGATARAATAAATAAATAAATAAATKNTTTTTTTTCTTT 691
Db 750 ATATAAATGTAKTGTAGAGATARAATAAATAAATAAATAAATKNTTTTTTTTCTTT 691

QY 127 CAACGGGTAAAGAGCGCTTGGTAAATTCGCATTA 161
Db 127 CAACGGGTAAAGAGCGCTTGGTAAATTCGCATTA 161

QY 690 AAWTAGGTATAAAAGAGGTGKTTTGTGTAATTA 656
Db 690 AAWTAGGTATAAAAGAGGTGKTTTGTGTAATTA 656

RESULT 10
BX447889      1186 bp    mRNA    linear    EST 22-MAY-2003
LOCUS         BX447889 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
DEFINITION    clone CS0DH08YD21 5-PRIME, mRNA sequence.
ACCESSION     BX447889
VERSION       BX447889.1 GI:31025888
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               1 (bases 1 to 1186)
               Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
               Full-length cDNA libraries and normalization
               Unpublished
               Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :
               http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
               Faraday Avenue Genoscope sequence ID : GSOAH008CB11Qp1.
               Location/Qualifiers
               1..1186
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DH08YD21"
               /tissue_type="T CELLS (JURKAT CELL LINE)"
               /cell_line="JURKAT CELL LINE"
               /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
               /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
               with a NotI-oligo(dT) primer. Five prime end enriched,
               double-strand cDNA was digested with Not I and cloned into
               the Not I and EcoRV sites of the pCMVSPORT 6 vector.
               Library was not normalized."
BASE COUNT    298 a 273 c 285 g 185 t 145 others
ORIGIN

```

```

Query Match      13.8%; Score 37; DB 13; Length 1186;
Best Local Similarity 36.8%; Pred. No. 22; Mismatches 58; Indels 0; Gaps 0;
Matches 53; Conservative 33;

QY 70 TGAGATCGGTGGAACAGAGGGAAGAAACAAAGCCTCCCGTGGGAAGAAATCCAA 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 881 KKAARWGGGCCCAADDTATGKATWAAAMVAASCMYCCSGKKGKGAATAATTTAA 940
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 CGGTAAAAAGCCCTGGTAATGTCATATCGGCTTGCAACACAAATTCGAATT 189
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 941 RGTWARTGNVMSKGRKTKGTGGAARAMCCMSSGKTKGGGAAAATTHMAWCC 1000
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 CTACTATCCAGGAAACGAAAGG 213
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1001 CCWAATGGRABCCCVGAAGGR 1024
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
AW942932/c      521 bp mRNA linear EST 23-APR-2001
LOCUS LD29090.3prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD29090 3, mRNA sequence.
ACCESSION AW942932
VERSION AW942932.1 GI:8120637
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 521)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished
COMMENT Other ESTs: LD29090.5prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AB003678: Drosophila melanogaster genomic scaffold
142000013386035 section 3 of 105, complete sequence.: 02/25/2001
Plate: LD 290 row: H column: 6
High quality sequence stop: 382.
Location/Qualifiers
1..521
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LD29090"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XLI Blue"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/notes="Organ: embryo; Vector: pOT2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."

BASE COUNT 147 a 129 c 100 g 145 t
ORIGIN
Query Match      13.7%; Score 36.8; DB 10; Length 521;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CGGAGCATAAATCGTTTAAAGATAATGTTATAGGAACAGCTTCGATGAAATCAATG 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 CCGAGCATCAATCGTTTAAAGATAATGTTATAGGAATAATTTGTTTAATATATG 190
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 61 ATTTTATTGAGAAATCGTCCGAAACAGAGGGAAG 96
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 ATATATTTTCTAGGTATACGTTTACAGAGTGCGATG 154
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
AJ449657
LOCUS Gallus gallus bursal lymphocyte EST
DEFINITION AJ449657 riken1: Gallus gallus cDNA clone 22ml2r1, mRNA sequence.
ACCESSION AJ449657
VERSION AJ449657.1 GI:20216878
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 763)
AUTHORS Buerstedde,J.M.
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..763
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="22ml2r1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/clone_lib="riken1"
/notes="CB inbred strain"

BASE COUNT 231 a 163 c 209 g 158 t
ORIGIN
Query Match      13.6%; Score 36.6; DB 9; Length 763;
Best Local Similarity 51.2%; Pred. No. 27;
Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 7 ATAAATCGTTTATTAAAGATATGTTAGGAACGAGTTCGATCAAAATCAATGATTTT 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 512 ATCAGACAGTTAAAGAGACCATGCAAAATGAACTGCTCAGATGAGGTACGATTATT 571
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 ATTTGAGAAATCGTCCGAAACAGAGGGAAGAAACAAAGCCTCCCGTCCGGAAGATTC 126
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 572 CTTTCAGTAAAGGAGGGGAAAAAACTGAAAGAGAGCTCAAGCCCACTGATTAAGATTA 631
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 CAACGGTAAAAAAGCCCTTCGTAAATTTGGATTATTCGGGCTT 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 632 TTGAGATGAGATTTCTTGATCAGTTNGAATTTGTTGCTT 675
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
AL534914/c
LOCUS AL534914 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
DEFINITION CSDF0007YCI3 3-PRIME, mRNA sequence.
ACCESSION AL534914
VERSION AL534914.2 GI:30541177
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished

```

On Feb 13, 2001 this sequence version replaced GI:12798407.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4660.f For
more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF007AB07NP1&cluster=4660.f)
cgi-bin/cluster.cgi?seq=CS0DF007AB07NP1&cluster=4660.f. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF007AB07NP1.

```

FEATURES
source
    Location/Qualifiers
        i. 1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DF007YC13"
            /tissue_type="FETAL BRAIN"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL BRAIN"
            /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-strand cDNA was digested with NotI and
                cloned into the NotI and EcoRV sites of the pCMVSPORT 6
                vector. Library was not normalized."
            285 a      221 c      288 g      344 t      63 others
BASE COUNT

```

[illegible]

RESULT 14
CNS0152W
LOCUS
DEFINITION
1310 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN15124 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
GSS.
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1310)

TITLE
 Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : secre@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Direct submission
 COMMENT
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) at
<http://www.edgp.ebi.ac.uk> . This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaut at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton

```

and Genevieve Payan. It has been constructed in the vector
pBelOAc11.
FEATURES             Location/Qualifiers
     source           1..1310
                        /organism="Drosophila melanogaster"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:7227"
                        /clone="BACN15124"
                        /clone_lib="DrosBAC"
                        /plasmid="pBelOAc11"
                        /note="end : SP6"
BASE COUNT          330 a   140 c   396 t   225 others
ORIGIN

```

[illegible]

RESULT 15	
AZ674925/c	
LOCUS	885 bp DNA linear GSS 14-DEC-2000
DEFINITION	ENTJW21fr Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION	AZ674925
VERSION	AZ674925
KEYWORDS	GI:11812071
SOURCE	GSS.
ORGANISM	Entamoeba histolytica
REFERENCE	Entamoeba histolytica
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.
TITLE	1 (bases 1 to 885)
JOURNAL	Loftus, B., Van Aken, S. and Praeger, C.
COMMENT	Determination of clone end sequences from Entamoeba histolytica
	HM1:IMSS sheared DNA library
	Unpublished
	Contact: Brendan J. Loftus

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftust@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 19
High quality sequence stop: 853.

```

FEATURES
source
high quality sequence stop: 699.
Location/Qualifiers
1..885
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba
/notes=Vector: pHOs1;
Institute for Genomic Research (TIGR), Rockville, MD.
histolytica Sheared DNA"
Site 1: Bst I; Constructed at The

```


RESULT 2

NAME/KEY: misc.feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc.feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 12.0%; Score 32.4; DB 4; Length 1664976;
Best Local Similarity 50.6%; Pred. No. 11;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 56 CAATGATTTTATTTAGAGATCGTTCGGAACAGAGGGAAGAAAAACAAGCCTCCCGTG 115
Db 1414104 CTAUGTTTTCGGCCAGAGATGATAATATATAATGGAACACAGGCAATCCAGG 1414163

Qy 116 CGGAAGAATTCACGGGTAAAAAGCCCTTGTAAATGTCATTTTCGGCTTGGCAA 175
Db 1414164 GCATAACATTCATAGCTGTTTTCGGAGCTTTTATTTTGGCATTTGATGTCAGGATT 1414223

Qy 176 CAATATTCGAATTTCTACTATCCAGGAAACGAA 209
Db 1414224 CAACATTGGAAGTGTCTTCGGCATAGCTAATGGA 1414257

RESULT 3
US-09-056-075-1/c
; Sequence 1, Application US/09056075

Patent No. 5955368
GENERAL INFORMATION:
APPLICANT: Johnson, Eric A.
APPLICANT: Bradshaw, Marite
APPLICANT: Rood, Julian
TITLE OF INVENTION: Expression System for Clostridium
NUMBER OF INVENTION: Species
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,075
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95238
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3770..4013
OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
plasmid RP4"
US-09-056-075-1

Query Match 11.9%; Score 32; DB 2; Length 6243;
Best Local Similarity 48.4%; Pred. No. 2.7;
Matches 89; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 8 TAAATCGTTATTAAAGATAATCGTATAGAACGAGCTTCGATGAATCAATGATTTTA 67
Db 1771 TTATAGATTTATTAGAAATTCAGAGAAGAACAAAAATATATGACAAATAATTTCTA 1712

Qy 68 TTTGAGATCGTTCGGAACAGAGGGAAGAAAAACAAGCCTCCCGTGGGAGAAATTC 127
Db 1711 AAAAGGAATATAGAGAAGAGAAAAATATTAGAGGTAAAAAAATATCAAGAGCAATTAA 1652

Qy 128 AACGGGTAAAAAGCGCTTGGTAAATTTGCGGCTTGGCAACAAAATTCGAAT 187
Db 1651 AACCTAAGGAAAGCAACAAAAAGAGGAATTAATGTTATTAGAAAAAAATAAAG 1592

Qy 188 TTCT 191
Db 1591 CCCT 1588

RESULT 4
US-09-134-001C-123
; Sequence 123, Application US/09134001C
; Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 123

; LENGTH: 435

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-123

Query Match 11.8%; Score 31.8; DB 4; Length 435;

Best Local Similarity 53.7%; Pred. No. 1.4;

Matches 66; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 9 AAATCGTTATTAAAGATAATGGTATAGCAAGAGCTTCGATGAAATCAATGATTTTAT 68

Db 192 AGATTACTATTAGATATGAGCTTTGTTGGTGATGATGATGAAAGCATTTGAAACCTAC 251

QY 69 TTGAGAAATCGGTCGGAACAGAGGAAGAAACAAAGCCTCCCGTCGGAAGAAATTCGA 128

Db 252 AAGAGTAATATGAATACTAGGTATATAAATACACTTTCTTCATGCACATTTATTCGA 311

QY 129 ACG 131

Db 312 AAG 314

RESULT 5

US-09-107-532A-2727/c

; Sequence 2727, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Arinello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 2727:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1167 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (b) LOCATION 1...1167

; SEQUENCE DESCRIPTION: SEQ ID NO: 2727:

US-09-107-532A-2727

Query Match 11.7%; Score 31.6; DB 4; Length 1167;

Best Local Similarity 51.4%; Pred. No. 2.2;

Matches 73; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 10 AAATCGTTATTAAAGATAATGGTATAGCAAGAGCTTCGATGAAATCAATGATTTTAT 69

Db 1064 AAATCAGTGTAGAAATAAAGTTTATACAAACTTTTCGCTGATCAAGTTGACGAGATTTAAT 1005

QY 70 TCAGAAATCGTCGGAACAGAGGGAAGAAACAAAGCCTCCCGTCGGAAGAAATTCGA 129

Db 1004 TCACCAACAATCGGAATATGAAGATCAATCCAATCTTCACTTGTTCGCCCAATCAT 945

QY 130 CGGGTAAAAAAGGCCCTTGGTAA 151

Db 944 CCAGCAAAAAGCTCATCGCA 923

RESULT 6

US-09-134-001C-135

; Sequence 135, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 135

; LENGTH: 1713

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-135

Query Match 11.5%; Score 31; DB 4; Length 1713;

Best Local Similarity 51.9%; Pred. No. 3.7;

Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 97 AAAACAAGCTCCCGTCGGAAGAAATCCAAAGGTTAAAAAGGCCCTTGTAAATTGG 156

Db 1284 AAAACAACAAAAAATAATGAAGAATTAATCTGGGAATATTTTAATCCCAATTA 1343

QY 157 CATTATTCGGGCTTGGCAACAAAATTGAAATTTCTACTATCCAGGAAACGAAAGGTTA 216

Db 1344 CAAAAGCTTTCAGAACCCAGAAACAAAAAATATCTACAGTCAAGAAAAATTGATGGA 1403

QY 217 CTTTACITTTGGATT 231

Db 1404 ACCTGTTTTTGGATT 1418

RESULT 7

US-09-134-001C-2824

; Sequence 2824, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

```
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2824
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2824

Query Match      11.5%; Score 31; DB 4; Length 1920;
Best Local Similarity 46.9%; Pred. No. 3.8;
Matches 97; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY      20 TAAAGATATGTTATAGAACAGCTTCGATGAATCAATGATTTTATTTGAGATCGG 79
Db      1341 TGATTTGATCGTGTAGACCCGACACTTATTTTAAATGATGAAGATTATCAAAAAATTAG 1400

QY      80 TCGGAACAGAGGGAAGAAACAAAGCCTCCCGTGGCGAAGATTCACACGGGTAAAA 139
Db      1401 AGACATATRAAGAAAGAAATATGTTTCCCAATATGTTTGAATCTTAAAAATAAAAA 1460

QY      140 AGGCTTGGTAAATGGCATTATTCGGCTTGGCAACAAAATGAAATTTCTACTATCCA 199
Db      1461 TGATTTACCAATATAGAAAAATAGTTAGTAGTATCAATCAAGATATAGAAACAAGAAG 1520

QY      200 GGGAAACGAAGGGTACTTTACTTT 226
Db      1521 TGACCGCAAGTGAATTTCTAGTTT 1547

RESULT 8
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..<15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..<30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..<45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..<60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..<75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..<90000)
; OTHER INFORMATION: n=a or c or g or t

; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..<105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..<120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..<135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..<150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..<165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..<180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..<195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..<210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..<225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..<240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..<255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..<270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..<285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..<300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..<315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..<330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001)..<345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001)..<360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001)..<375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001)..<390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001)..<405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001)..<420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001)..<435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001)..<450000)
; OTHER INFORMATION: n=a or c or g or t
```

NAME/KEY: misc feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

Query Match 11.5%; Score 31; DB 4; Length 1230025;
Best Local Similarity 64.8%; Pred. No. 25;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 175 ACAAAATTCGAATTTCTACTATCCAGGAAACGAAAGGGTTACTTTTACTTTGATTATG 234
Db 797606 ACCAACTACAAATCTATTATCCAGAAACAGGACTGTACGCAAAATATCGATTATG 797547
QY 235 CAGAGATTGCA 245
Db 797546 AAGATATTGGA 797536

RESULT 9

US-08-181-271A-106/c
Sequence 106, Application US/08181271A
Patent No. 5614395

GENERAL INFORMATION:

APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.

TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne

STATE: New York

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/181,271A
; FILING DATE: 13-JAN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-181-271A-106

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5;
Matches 110; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY - 21 AAAGATAATGGTATAGGAACGAGGTTCCGATGAATCAATGATTTTATTGAGATCGGT 80
DB 256 AAACATTTGCTTTGTTTGGTTTTCACATGACATCGAGTTGCTGACACGACGAGGT 197
QY 81 CGGAACAGAGGGGAGAAAACAGACCTCCCGTGGGAGAAATCCCAACGGGTAAAAA 140
DB 196 CAGTGTACCGCGCGAGTAATCACCAGTACCCATGGAACAAATGGAGACCTGGATTAA 137
```

```

QY 141 GGCTTTGGTAATGGCATTTATTCGGGCTTGGCAACAAAATTTCTACTATCCAG 200
DB 136 TACTTTTGGAAATTCGCAAAATCTCTCCCTTGACCGCAACATCTGATTGCTTCAAGCAC 77
QY 201 GGAACGAAAGGGTACTTTTACTTTTGGATTATCGAGAGATTTCGAGACGACGAGGTATT 260
DB 76 GCACACGAAAAGATTCCCATTTATTCGACAAAGCAAGAGATTATGATGGCTTAT 17
QY 261 TAT 263
DB 16 TAT 14

RESULT 10
US-08-449-315-106/c
; Sequence 106, Application US/08449315
; Patent No. 5650505
; GENERAL INFORMATION:
; APPLICANT: Rvals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr. Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,315
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
```

FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8699
INFORMATION FOR SEQ ID NO: 106:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-449-315-106

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5;
Matches 110; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 21 AAAGATATGGTATAGGAACGAGCTTCGATGAATCAATGATTTTATTTGAGATCGGT 80
DB 256 AAACATATGCTTTGTTTGGTTTTCACATGACATCGAGTTGCTGAAGACGAGGT 197

QY 81 CGAACAAGAGGAG 140
DB 196 CAGTGACGGGCGAGCTAATCAGCATGACCCATGGAACAAATGGAGACCTGGATTAA 137

QY 141 GCGCTTGTAAATTCGATTTTCGGCTTGGCAACAAATGAAATTTCTACTATCCAG 200
DB 136 TACTTTTGGAAATTTGCAAAATCTCTCTTACCGCAACATCTGATTCCTTCAACGAC 77

QY 201 GAAACAGAAAGAGGTTACTTTTACTTTGATTTGATTCAGAGATTCGAAGAGAGAGGTTAT 260
DB 76 GCAACAGAAAGATTCCTCCCAATTAATCTGAACAGAGCAAGATTAATGAATGCTTAT 17

QY 261 TAT 263
DB 16 TAT 14

RESULT 11
US-08-444-803-106/C
Sequence 106, Application US/08444803
Patent No. 5654414
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.

APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,803
FILING DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-444-803-106

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5;
Matches 110; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 21 AAAGATATGGTATAGGACGAGCTTCGATGAATCAATGATTTTATTGAGATCGGT 80
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 AAACACATTGCTTTGTTTGGTTTCACATCGACATCGAGTTGCTGAAGACGAGGTT 197
QY 81 CGGACAGAGGAGGAGAGAAACAAAGCCTCCCGTCGGGAGAAATTCGACGGGTAAAAA 140
DB 196 CAGTGTACGGGACGTAATCACCATGACCCATGGAACAAATGGAGACCTGGATTAAG 137
QY 141 GGCCTTGTAATTCGGCATTTATTCGGGTTGGCAACAAAATTTGAATTTTACTATCCAG 200
DB 136 TACTTTTGGAAATTTGCCAAATCTCTCTTGACCGCAACATCTGATTCCTTCACAAGCAC 77
QY 201 GCAACGAGAGGTTACTTTTACTTTGATTTGATGAGAGATTCGAGAGCAAGGTTATT 260
DB 76 GCAACGAGAGATTTCCCATTTTCTGAAAGAGCAAGCAAGATTTAATGATGCTTAT 17
QY 261 TAT 263
DB 16 TAT 14

RESULT 12
US-08-043-106/c
; Sequence 106, Application US/08449043
; Patent No. 5689044
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Unnes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; US-08-449-043-106

STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,043
FILING DATE: 24-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-449-043-106

```

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5; Mismatches 133; Indels 0; Gaps 0;
Matches 110; Conservative 0;

21 AAAGATAATGGTATAGGAACGAGCTTCGATGAAATCAATGATTTTATTTGAGAAATCGGT 80
|||
256 AACACATTTGCTTTGTTTTCACATGACATCGAGTTGCTGAACGAGGTT 197
|||
81 CGAACAGAGGAGGAGAAACAAACGCTCCCGTCGGAAGAAATCCAAACGGGTAAAAA 140
|||
196 CAGTGTACCGGCGAGCTTAATCACCAGTACCCATGGAACAAATGGAGACCTGGATTAA 137
|||
141 GGCCTTGGTAAATTCGCATTATTCGGCTTGCAACAAATTTGAATTTCTACTATCCAG 200
|||
136 TACTTTTGGAAATTTGCCAATCTCTCTTGACCCACATCTGTTGCTTCACAGCAC 77
|||
201 GGAACAGAAAGGTTACTTTTACTTTTGGATTATCGAGAGATTCGAGAGCAAGGTTAT 260
|||
76 GCAACAGAAAGATTTCCCAATTTCTCTGAAACAGAGCAAGATTATGAATGCTTAT 17
|||

261 TAT 263
16 TAT 14

RESULT 13
US-08-456-265A-106/c
Sequence 106 Application US/08456265A
Patent No. 5767369
GENERAL INFORMATION:
APPLICANT: Alexander, Danny C.
APPLICANT: Ryals, John A.
APPLICANT: Goodman, Robert M.
APPLICANT: Stinson, Jeffrey R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESS: CIBA-GEIGY Corporation
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,265A
FILING DATE: 31-MAY-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/181,271
FILING DATE: 13-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727/DIV10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-456-265A-106

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5; Mismatches 133; Indels 0; Gaps 0;
Matches 110; Conservative 0;

21 AAAGATAATGGTATAGGAACGAGCTTCGATGAAATCAATGATTTTATTTGAGAAATCGGT 80
|||
256 AACACATTTGCTTTGTTTTCACATGACATCGAGTTGCTGAACGAGGTT 197
|||
81 CGAACAGAGGAGGAGAAACAAACGCTCCCGTCGGAAGAAATCCAAACGGGTAAAAA 140
|||
196 CAGTGTACCGGCGAGCTTAATCACCAGTACCCATGGAACAAATGGAGACCTGGATTAA 137
|||
141 GGCCTTGGTAAATTCGCATTATTCGGCTTGCAACAAATTTGAATTTCTACTATCCAG 200
|||
136 TACTTTTGGAAATTTGCCAATCTCTCTTGACCCACATCTGTTGCTTCACAGCAC 77
|||
201 GGAACAGAAAGGTTACTTTTACTTTTGGATTATCGAGAGATTCGAGAGCAAGGTTAT 260
|||
76 GCAACAGAAAGATTTCCCAATTTCTCTGAAACAGAGCAAGATTATGAATGCTTAT 17
|||

261 TAT 263
16 TAT 14

RESULT 14
US-08-456-416-106/c
Sequence 106 Application US/08455416
Patent No. 5777200
GENERAL INFORMATION:
APPLICANT: Ryals, John A.

APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSES: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.125
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,416
FILING DATE: 31-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-455-416-106
Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred.No. 5.5;
Matches 110; Conservative 0; Mismatches 133; Indels 0; Caps 0;
QY 21 AAAGATAATGGTATAGGACGAGCTTCGATGAAATCAATGATTTTATTGAGAAATCGGT 80
Db 256 AAACACATTGCTCTTTGTTTGTTCACATGACATCGAGTTGCTGAAGACGAGGGT 197
QY 81 CGGAACAGAGGGAAGAAACAAAGCCTCCCGTGGGAGAAATTCACACGGGTAAAAA 140
Db 196 CAGTGACCGGCGACGTAATACCCGATGACCAATGGAACAAATTTGGAGACCTGGATTAG 137
QY 141 GGCTTTGGTAAATGGCATTATTCGGGCTTGGCAACAAAATTTGAAATTTCTACTATCCAG 200
Db 136 TACTTTTGGATTTGCCAAATCTCTCTTGACCGCAACATCTGATGCTTCAACAGCAC 77
QY 201 GGAACGAAAGGTTACTTTTACTTTTGGATTATGACAGATTGGAAGAGCAAGGTTAT 260
Db 76 GCAACAGAAAAGATTCCTCCATTATTCTGAACAGCAAGCAAGAGATTATGATGATGCTTAT 17
QY 261 TAT 263
Db 16 TAT 14
RESULT 15
US-08-455-244-106/c
Sequence 106, Application US/08455244
Patent No. 5789214
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,244
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmet, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/F1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-455-244-106

Query Match 11.2%; Score 30.2; DB 1; Length 1031;
Best Local Similarity 45.3%; Pred. No. 5.5; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 133;
QY 21 AAAGATAATGGTATAGGAACGAGCTTCGATGAAATCAATGATTTTATTTGAGATCGGT 80
Db 256 AAACACATTGCTCTTGTGTTTTCACACTGAACATCGAGTTGCTGAAGACGAGGGT 197
QY 81 CGAACAGAGGGAAGAAAAACAGCCCTCCCGTCCGGAAGATTCACAGCGGTAAAAA 140
Db 196 CAGTGTACCGGCGACGTAAATCACCAGATGACCCATGGAAACAAATTTGGAGACTGGATTAG 137
QY 141 GGCCTTGGTAAATTTGGCATTATTGGGCTTGGCAACAAAAATTTGAAATTTCTACTATCCAG 200
Db 136 TACTTTTGGAAATTTGCCAAATCTCTCTTGACCGCAACATCTGATTGCTTCAAGCAC 77
QY 201 GGAACAGAGGAGGTACTTTTACTTTGGATTATCGAGAGATTGGAAGAACGAGGGTATT 260
Db 76 GCACAGAAAAAGATTCCCCATTATTCTGAACAGAGAAAGAGATTATGATGGCTTAT 17
QY 261 TAT 263
Db 16 TAT 14

Search completed: November 15, 2003, 08:09:05
Job time : 46.8989 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 152.329 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-77
Perfect score: 269
Sequence: 1 CGGAGCATAAATCGTTATT.....GCAGGGTATTATCAACCG 269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	269	100.0	269	10	US-09-928-457-77		Sequence 77, Appl
2	121.8	45.3	270	10	US-09-928-457-77		Sequence 88, Appl
3	37.2	13.8	914	12	US-10-027-632-153460		Sequence 153460,
4	37.2	13.8	914	13	US-10-027-632-153460		Sequence 153460,
5	36.8	13.7	1596	14	US-10-198-846-13583		Sequence 13583, A
6	36	13.4	2000	10	US-09-938-842A-4782		Sequence 4782, Ap
7	36	13.4	185695	14	US-10-020-141-11		Sequence 11, Appl
8	36	13.4	185695	14	US-10-017-721-1		Sequence 11, Appl
9	33	12.3	1142	13	US-10-001-876-2		Sequence 2, Appl1
10	32.6	12.1	557	12	US-10-027-632-138807		Sequence 138807,
11	32.6	12.1	557	12	US-10-027-632-138808		Sequence 138808,
12	32.6	12.1	557	13	US-10-027-632-138807		Sequence 138807,
13	32.6	12.1	557	13	US-10-027-632-138808		Sequence 138808,
14	31.8	11.8	6118	12	US-10-027-632-138808		Sequence 1004, Ap
15	31.6	11.7	604	12	US-10-027-632-72923		Sequence 72923, A
16	31.6	11.7	604	12	US-10-027-632-74019		Sequence 74019, A

c 17	31.6	11.7	604	13	US-10-027-632-72923	Sequence 72923, A
c 18	31.6	11.7	604	13	US-10-027-632-74019	Sequence 74019, A
19	31.6	11.7	3673778	12	US-10-312-841-1	Sequence 1, Appli
20	31.4	11.7	884	14	US-10-106-698-356	Sequence 356, App
21	31.4	11.7	2200	12	US-10-311-455-2164	Sequence 2164, Ap
c 22	31.2	11.6	613	12	US-10-027-632-174871	Sequence 174871,
c 23	31.2	11.6	613	13	US-10-027-632-174871	Sequence 174871,
24	31	11.5	2106	12	US-10-253-157-375	Sequence 375, App
25	31	11.5	4348	12	US-09-814-353-22000	Sequence 22000, A
26	31	11.5	5222	12	US-10-006-285-443	Sequence 443, App
27	31	11.5	186957	14	US-10-185-770-3	Sequence 3, Appli
28	30.8	11.4	420	12	US-09-814-353-13772	Sequence 12772, A
29	30.8	11.4	421	12	US-09-814-353-5169	Sequence 5169, Ap
30	30.8	11.4	1989	10	US-09-070-927A-440	Sequence 240, App
31	30.6	11.4	598	12	US-10-027-632-238992	Sequence 238992,
32	30.6	11.4	598	13	US-10-027-632-238992	Sequence 238992,
c 33	30.6	11.4	637	12	US-10-027-632-235142	Sequence 235142,
c 34	30.6	11.4	637	13	US-10-027-632-235142	Sequence 235142,
c 35	30.6	11.4	660	14	US-10-106-698-586	Sequence 586, App
c 36	30.6	11.4	32768	10	US-09-070-927A-17	Sequence 17, Appl
c 37	30.2	11.2	497	11	US-09-764-891-6683	Sequence 6683, Ap
c 38	30.2	11.2	497	14	US-10-091-572-544	Sequence 544, App
39	30.2	11.2	1314	12	US-10-027-632-254785	Sequence 254785,
40	30.2	11.2	1314	13	US-10-027-632-254785	Sequence 254785,
c 41	30	11.2	815	12	US-10-027-632-154727	Sequence 154727,
c 42	30	11.2	815	12	US-10-027-632-154728	Sequence 154728,
c 43	30	11.2	815	13	US-10-027-632-154727	Sequence 154727,
c 44	30	11.2	815	13	US-10-027-632-154728	Sequence 154728,
c 45	30	11.2	6059	12	US-10-311-455-1453	Sequence 1453, Ap

ALIGNMENTS

RESULT 1

US-09-928-457-77
; Sequence 77, Application US/09928457
; Patent No. US20020164603A1

GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: DNA, specific proteins and peptides

; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method

; TITLE OF INVENTION: for obtaining them and their biological application.

; NUMBER OF SEQUENCES: 99

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEB)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/928,457

; FILING DATE: 2001-08-14

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/214,759

; FILING DATE: 199-12-10

; INFORMATION FOR SEQ ID NO: 77:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 269 base pairs

; TYPE: nucleotide

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-09-928-457-77

Query Match 100.0%; Score 269; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 5.5e-68;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGACGAGCTTCGATGAATCAATG 60

Db 1 CGGAGCATAAATCGTTATTAAAGATAATGGTATAGGACGAGCTTCGATGAATCAATG 60

```

QY 61 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAAACAAAGCCTCCCGTGGCGAA 120
Db 61 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAAACAAAGCCTCCCGTGGCGAA 120
QY 121 GAATTCCAACGGGTAATAAAGCCCTTGTAATTTGGCAATATTCGGGCTTGGCAACAAA 180
Db 121 GAATTCCAACGGGTAATAAAGCCCTTGTAATTTGGCAATATTCGGGCTTGGCAACAAA 180
QY 191 TTGAATTTCTACTATCAGGGAACGAAGGTTACTTTTACTTTTGGATTATGCAGAGA 240
Db 181 TTGAATTTCTACTATCAGGGAACGAAGGTTACTTTTACTTTTGGATTATGCAGAGA 240
QY 241 TTCGAAGAAGCAAGGTTATTTATCAACCG 269
Db 241 TTCGAAGAAGCAAGGTTATTTATCAACCG 269

```

RESULT 2

```

US-09-928-457-88/c
; Sequence 88, Application US/09928457
; Patent No. US2002016403A1
; GENERAL INFORMATION:

```

```

; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (OEB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-928-457-88

```

```

Query Match 45.3%; Score 121.8; DB 10; Length 270;
Best Local Similarity 98.4%; Pred. No. 2.3e-25;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGACATAAAATCGTTATTAAAGATAATGTTATAGGAACGAGCTTCGATGAAATCAATG 60
Db 125 CGGACATAAAATCGTTATTAAAGATAATGTTATAGGAATGAATTCGATGAAATCAATG 66
QY 61 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAAACAAAGCCTCCCGTGGCGAA 120
Db 65 ATTTTATTTCAGAAATCGTTCGGAACAGAGGGAAGAAAACAAAGCCTCCCGTGGCGAA 6
QY 121 GAATT 125
Db 5 GAATT 1

```

RESULT 3

```

US-10-027-632-153460
; Sequence 153460, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```

```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 32520
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153460
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-153460

```

```

Query Match 13.8%; Score 37.2; DB 12; Length 914;
Best Local Similarity 50.6%; Pred. No. 1.1;
Matches 87; Conservative 1; Mismatches 84; Indels 0; Gaps 0;

```

```

QY 92 GCAGAGAAAACAGGCTCCCGTGGGAGAGATTCCAACGGGTAAAAAGCGCTTGTTAA 151
Db 137 GCAGAGAAAACATGATGATAGGCAAGAGATTTGGAATGATTAATTAAGTTAC 196
QY 152 ATTGCATTATTTCGGGCTTGGCAACAAAATTGAAATTTCTACTATCCAGGAAACGAAAG 211
Db 197 TAATAATATATTAGTAATTTCAACAAAAGAGATTTTAAGCTATAAGTTCAAAAACATAG 256
QY 212 GGTACTTTTACTTTGGATTATGCGAGATTCGAGAGATTCGAGAGACGAGGTTATAT 263
Db 257 GGAAGATTTTGGGAGAACTACCAATTTTCCAGAAAACAAATATTTGT 308

```

RESULT 4

```

US-10-027-632-153460
; Sequence 153460, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 32520
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153460
; LENGTH: 914
; TYPE: DNA

```

```

; ORGANISM: Human
US-10-027-632-153460

Query Match      13.8%; Score 37.2; DB 13; Length 914;
Best Local Similarity 50.6%; Pred. No. 1.1;
Matches      87; Conservative 1; Mismatches 84; Indels 0; Gaps 0;

QY  92  GGAAGAAAACAGCCCTCCCGTCGGAGGATTCACAGCGGTAAAGAGCCCTTGCTAA 151
DB  137  GCAAGAAAACATGAAGTAGGCCAAGAAATTTGGAAATGCAATTAATTAAGTAGTTAC 196

QY  152  ATTGGCATTATTGGCGCTTGGCAACAAATTTGAAATTTCTACTATCCAGGSAACGAAAG 211
DB  197  TAATAATATATTAGTAATTRCAACAAACAAAGAGATTTAAGCTATAAGTTCCAAACATAG 256

QY  212  GGTACTTTTACTTTGGATTATGAGAGATTCGAGAGCAAGCGTATTAT 263
DB  257  GGAAGATTTTAGGAGAACTACCAAAATTTTCCAGGAAACAAATATTGT 308

RESULT 5
US-10-198-846-13583
; Sequence 13583, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yongxiao
; APPLICANT: Lillie, James
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13583
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 1484, 1497,
; LOCATION: 1554, 1578, 1585, 1591, 1596
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13583

Query Match      13.7%; Score 36.8; DB 14; Length 1596;
Best Local Similarity 48.6%; Pred. No. 1.9;
Matches 101; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY  53  AATCAATCATTTTATTGAGATCGGTGCGAAGGAGGAAAGAAAACAGCCTCCCC 112
DB  1138  AAACCTCGTTTGTAGTAAAGAGGCTGGGAAGAGGAGGAGCCACAAATCTGTCTGC 1197

QY  113  GTGCGGAAGAATTCCACGGGTAAAGAGCCCTTGGTAAATTTGGCAATTTTCGGCCTGG 172
DB  1198  TTCTCACAATAGTCATTGGGCAATAAGGCATCTGTCTTTTGGGCTGCTGCCITCA 1257

QY  173  CAACAAATTTGAATTTCTACTATCCAGGAAACGAAGGTTACTTTTACTTTGGATTA 232
DB  1256  GCACAGAGGAGCCGCAAACTCTATTTCGGGGCAACGAGGGTTAACCATCTCTTCAGTTGA 1317

QY  233  TCCAGAGATTGGAAGAACGAGGTTATT 260
DB  1318  ACCAGAGTTGACAAGCCCTATGGGAAT 1345

RESULT 6
US-09-938-842A-4782/c
; Sequence 4782, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepis, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4782
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4782

Query Match      13.4%; Score 36; DB 10; Length 2000;
Best Local Similarity 55.6%; Pred. No. 3.5;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY  118  GAAGATTTCCACGGGTAAAGAGCCCTTGGTAAATTTGCATTTTCGGGCTTGGCAACA 177
DB  767  GAACAAATAAAGAGTAAAGATTTTCATGGTAAATAGGAAATCTATTTTATTTAAAA 708

QY  178  AAATGAAATTTCTACTATCCAGGGAACGAAAGGTTACTTTTACTTTGGATTATGCAG 237
DB  707  AAATGAAATACATATCTTTATGAATTAATGAGAAATTTTAAATCGAATTCAGTTG 648

QY  238  AGAT 241
DB  647  AGTT 644

RESULT 7
US-10-020-141-11
; Sequence 11, Application US/10020141
; Publication No. US20030092013A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Ableson, Allen
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: NMI-002
; CURRENT APPLICATION NUMBER: US/10/020,141
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 185695
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-020-141-11

Query Match      13.4%; Score 36; DB 14; Length 185695;
Best Local Similarity 50.6%; Pred. No. 22;
Matches 87; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY  92  GGAAGAAAACAGCCCTCCCGTCGGAGGAAATTCACAGCGGTAAAGAGCCCTTGCTAA 151
DB  178160  GCAAGAAAACATGAAGTAGGCCAAGAAAGATTTGGAAATGCAATTAATTAAGTAGTTAC 178219
```


Search completed: November 15, 2003, 08:31:58
Job time : 156.429 secs

```
; LENGTH: 6118
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1004

Query Match          11.8%; Score 31.8; DB 12; Length 6118;
Best Local Similarity 53.7%; Pred. No. 91;
Matches 66; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 7 ATAAAAATCGTTATTAAAGATATGTTAGGACGAGCTTCGATGAATCAATCATTTT 66
DB 2177 AAAAAAATATAATAACAAAAAATATACGAAACTCGTCAATCTAATAATAACTTTA 2118

QY 67 ATTTGAGAAATCGTGGGACAGAGGGAAGAAAAACAAAGCTCCCGTGGCGAAGATTC 126
DB 2117 ATTAATAAATTCAAAAATATAAACAAATCATATAAAACCTTACCAATCAAAAAATCA 2058

QY 127 CAA 129
DB 2057 TAA 2055
```

```
RESULT 15
US-10-027-632-72923/c
; Sequence 72923, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72923
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72923

Query Match          11.7%; Score 31.6; DB 12; Length 604;
Best Local Similarity 53.2%; Pred. No. 40;
Matches 67; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 115 GCGGAAGAATTCACACGGTAAAAAGCCCTTGGTAATTCGCATTTATTCGGGCTTGCA 174
DB 532 GAGGCGAAGATCAATGCCVACTAGCTCTAGGAAATATGCGACATCTCAGAGAATAT 473

QY 175 ACAAAATTCGAATTTCTACTATCCAGGGAACGAAAGGTTACTTTTACTTTGGATTATG 234
DB 472 TTAATAATTTAAGGAATTAATAACAAATTAATAAACTTGATTTCTTTTGGAACTA 413

QY 235 CAGAGA 240
DB 412 CTGAAA 407
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 1013.71 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-79
Perfect score: 229
Sequence: 1 CGGTTTCAGGTTTCGCCAA.....TGGATGCAGCATCCATCCG 229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_ox.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	229	100.0	229	1	AF169453	AF169453 Neisseria
2	229	100.0	229	6	AE8908	AE8908 Sequence 79
3	229	100.0	229	6	BD063008	BD063008 DNA and s
4	229	100.0	349061	1	NMA224291	AL162753 Neisseria
5	227.4	99.3	15053	1	AE002550	AE002550 Neisseria
6	227.4	99.3	172325	6	AX044035	AX044035 Sequence
7	227.4	99.3	349980	6	AX044034	AX044034 Sequence
8	202.4	88.4	349	1	AF169440	AF169440 Neisseria
9	192.8	84.2	400	1	AF169418	AF169418 Neisseria
10	162.4	70.9	192	6	AE8896	AE8896 Sequence 67
11	162.4	70.9	192	6	BD062996	BD062996 DNA and s
12	161.2	70.4	409	1	NMAJ5415	AE005415 Neisseria
13	107	46.7	1275	6	AE6274	AE6274 Sequence 30
14	107	46.7	1572	6	AX044062	AX044062 Sequence
15	107	46.7	1575	6	AE6272	AE6272 Sequence 30
16	107	46.7	1576	6	AE6270	AE6270 Sequence 30
17	100.6	43.9	1575	6	AE6276	AE6276 Sequence 30
18	100.6	43.9	1575	6	AX024113	AX024113 Sequence
19	36.8	16.1	176638	2	AL590312	AL590312 Homo sapi
20	36.6	16.0	43345	9	AL590542	AL590542 Human DNA
21	35.2	15.4	156481	2	AC079838	AC079838 Homo sapi
22	35.2	15.4	157650	9	AC007098	AC007098 Homo sapi
23	35.2	15.4	167711	2	AC007629	AC007629 Homo sapi
24	35	15.3	13463	2	AC013907	AC013907 Drosophil
25	35	15.3	164969	3	AC023693	AC023693 Drosophil
26	35	15.3	179312	3	AC023745	AC023745 Drosophil
27	35	15.3	320855	10	AC125881	AC125881 Mus muscu
28	35	15.3	315790	3	AE003484	AE003484 Drosophil
29	35	15.3	334028	2	AC116537	AC116537 Drosophil
30	34.6	15.1	152572	9	AC090588	AC090588 Homo sapi
31	34.6	15.1	152881	2	AC025737	AC025737 Homo sapi
32	34.6	15.1	183169	9	AC009643	AC009643 Homo sapi
33	34.2	14.9	127705	5	AL929326	AL929326 Zebrafish
34	34.2	14.9	460920	8	AC137072	AC137072 Genomic s
35	34.2	14.9	203558	10	AL670464	AL670464 Mouse DNA
36	34	14.8	175524	2	AC140955	AC140955 Didelphis
37	33.8	14.8	99292	9	AL445216	AL445216 Human DNA
38	33.6	14.7	76013	2	AP005922	AP005922 Oryza sat
39	33.6	14.7	99200	2	AC136895	AC136895 Leishmani
40	33.6	14.7	112020	2	AL133285	AL133285 Homo sapi
41	33.6	14.7	113041	9	AL391297	AL391297 Human DNA
42	33.6	14.7	138485	9	AC105273	AC105273 Homo sapi
43	33.6	14.7	141791	2	AP005582	AP005582 Oryza sat
44	33.6	14.7	206030	9	AC021439	AC021439 Homo sapi
45	33.4	14.6	1622	9	BC050472	BC050472 Homo sapi

ALIGNMENTS

RESULT 1
AF169453
LOCUS AF169453 229 bp DNA linear BCT 09-AUG-2000
DEFINITION Neisseria meningitidis strain 22491 clone Cm043 unknown sequence.
ACCESSION AF169453
VERSION AF169453.1 GI:9754661
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria, Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 229)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to

JOURNAL	pathogenic Neisseriae
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 229)
TITLE	Perrin,A., Nassif,X. and Tinsley,C.R.
JOURNAL	Direct Submission
FEATURES	Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156 rue de Vaugirard, Paris 75015, France
source	Location/Qualifiers
BASE COUNT	1..229
ORIGIN	/organism="Neisseria meningitidis"
Query Match	100.0%; Score 229; DB 1; Length 229;
Best Local Similarity	100.0%; Pred. No. 5.6e-63;
Matches 229; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGTTTCAGGTTGTCGCGAAGGCTCGTAACGGCAACTGATTACGGGTGATCGAGCA 60
Db	1 CGTTTCAGGTTGTCGCGAAGGCTCGTAACGGCAACTGATTACGGGTGATCGAGCA 60
Qy	61 GCTTGACATTCGCGACGCGAGGCTGATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Db	61 GCTTGACATTCGCGACGCGAGGCTGATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Qy	121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAGACTATGCAAAA 180
Db	121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAGACTATGCAAAA 180
Qy	181 TATGAAAAACCAAGTACCGGATCAGGCATGGATGCACGATCCAATCCG 229
Db	181 TATGAAAAACCAAGTACCGGATCAGGCATGGATGCACGATCCAATCCG 229
RESULT 3	
BD063008	229 bp DNA linear PAT 27-AUG-2002
LOCUS	
DEFINITION	DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.
ACCESSION	BD063008
VERSION	BD063008.1 GI:22608611
KEYWORDS	JP 2001504684-A/70.
SOURCE	unidentified
ORGANISM	unidentified
REFERENCE	1 (bases 1 to 229)
AUTHORS	Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.
TITLE	DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications
JOURNAL	Patent: JP 2001504684-A 70 10-APR-2001; INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM
COMMENT	PN JP 2001504684-A/70 PD 10-APR-2001 PF 11-JUL-1997 JP 19980505685 PI 12-JUL-1996 FR 96/08768 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI CARLA VINALS, PI PETRA MERKER PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC Strandedness: Single; CC Topology: Linear; FH Key Location/Qualifiers.
FEATURES	Location/Qualifiers
source	1..229
BASE COUNT	72 a 48 c 65 g 44 t
ORIGIN	
Query Match	100.0%; Score 229; DB 6; Length 229;
Best Local Similarity	100.0%; Pred. No. 5.6e-63;
Matches 229; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGTTTTCAGGTTGTCGCGAAGGCTCGGTACGGGCAACCTGATTACGGGTGATCGAGCA 60
Db	1 CGTTTTCAGGTTGTCGCGAAGGCTCGGTACGGGCAACCTGATTACGGGTGATCGAGCA 60
Qy	61 GCTTGAACATTCGCGACGCGCAAGCGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Db	61 GCTTGAACATTCGCGACGCGCAAGCGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Qy	121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAGACTATGCAAAA 180
Db	121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAGACTATGCAAAA 180
Qy	181 TATGAAAAACCAAGTACCGGATCAGGCATGGATGCACGATCCAATCCG 229
Db	181 TATGAAAAACCAAGTACCGGATCAGGCATGGATGCACGATCCAATCCG 229
RESULT 2	
A68908	229 bp DNA linear PAT 06-MAY-1999
LOCUS	
DEFINITION	Sequence 79 from Patent WO9902547.
ACCESSION	A68908
VERSION	A68908.1 GI:4759827
KEYWORDS	unidentified
SOURCE	unidentified
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 229)
AUTHORS	Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and Merker,P.
TITLE	DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL APPLICATIONS
JOURNAL	Patent: WO 9802547-A 79 22-JAN-1998; INST NAT SANTE RECH MED (FR) Other publication FR 2751000 19980116.
COMMENT	Location/Qualifiers
FEATURES	1..229
source	/organism="unidentified"
BASE COUNT	72 a 48 c 65 g 44 t
ORIGIN	
Query Match	100.0%; Score 229; DB 6; Length 229;
Best Local Similarity	100.0%; Pred. No. 5.6e-63;
Matches 229; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGTTTTCAGGTTGTCGCGAAGGCTCGTAACGGCAACTGATTACGGGTGATCGAGCA 60
Db	1 CGTTTTCAGGTTGTCGCGAAGGCTCGTAACGGCAACTGATTACGGGTGATCGAGCA 60

```

RESULT 4
NMA222491/c
LOCUS
DEFINITION
  NMA222491 349061 bp DNA linear BCT 02-SEP-2002
  Neisseria meningitidis serogroup A strain Z2491 complete genome;
  segment 2/7.
ACCESSION
  AL162753 AL157959
VERSION
  AL162753.2 GI:7379120
KEYWORDS
  .
SOURCE
  Neisseria meningitidis Z2491
  Neisseria meningitidis Z2491
  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
  Neisseriaceae; Neisseria.
  1 (bases 1 to 349061)
REFERENCE
  Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
  Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
  Davies,R.M., Davis,P., Davis,K., Deakin,K., Feltwell,T., Hamlin,N.,
  Holtroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
  Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
  Skelton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
  Complete DNA sequence of a serogroup A strain of Neisseria
  meningitidis Z2491
  Nature 404 (6777), 502-506 (2000)
JOURNAL
  Nature 404 (6777), 502-506 (2000)
MEDLINE
  20222556
PUBMED
  10761919
REFERENCE
  2 (bases 1 to 349061)
AUTHORS
  Parkhill,J
DIRECT SUBMISSION
  Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
  sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
  Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
  Notes:
  Details of N. meningitidis sequencing at the Sanger Centre are
  available on the World Wide Web.
  (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
  source
    1..349061
    /organism="Neisseria meningitidis Z2491"
    /mol_type="genomic DNA"
    /strain="Z2491"
    /db_xref="taxon:122587"
    /note="serogroup: A"
    complement(24..206)
    /note="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
    310.39"
    /label=ATR
    209..212
    /gene="NMA0368"
    220..681
    /gene="NMA0368"
    220..681
    /gene="NMA0368"
    /note="NMA0368, probable integral membrane protein, len:
    153 aa; contains four probable transmembrane domains"
    /codon_start=1
    /transl_table=11
    /product="putative integral membrane protein"
    /protein_id="CAB83669.1"
    /db_xref="GI:7379121"
    /db_xref="SPTREMBL:Q9JWH7"
    /translation="MQEONRKSPPVIMLVLSVALIASLNVAFILGNHGSMEGLTV
    ILGSIFFASLDIRCAVYANYVMIAIIVLALRKXVPVHPAFWGLVAIFSVAIVV
    DEAGNTSDIVRYGAGFYLYWAAFAVASIGTFAGKXKKAASADGKNDV"
    671..1492
    /gene="hemK"
    671..1492
    /gene="hemK"
    /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
    to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
    scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
    PS00092 N-6 Adenine-specific DNA methylases signature"
    /codon_start=1
    /transl_table=11
  repeat_unit
    /product="HemK protein"
    /protein_id="CAB83670.1"
    /db_xref="GI:7379122"
    /db_xref="SPTREMBL:Q9JWH6"
    /translation="MTFDFKWLGLSKLPKNEARMLOVSEYTRVQLLTGRGEMPDVEV
    RQADRLAPRLNGEVAVILGARFVGRFTVNPVSLIPRPETELHVAVLARLPEN
    GRWVDTGSGAVAVTVALERPDFAFRASDISPDALETKNAADLCARVFAFGSWF
    DTDMPGSKWDIIVSNPPYIENGKHLSDGLRFEFCIALTDFSDLSLCIRTLQAQAP
    DRLAEGGFLLEHGFQGAARVGLAENGFSGVETLPDLAAGLDRVLTGLKYMHLK"
    831..840
    /gene="hemK"
    /note="Core DNA uptake sequence: gccgtctgaa"
    /label=DUS
    1162..1171
    /gene="hemK"
    /note="Core DNA uptake sequence: gccgtctgaa"
    /label=DUS
    1187..1207
    /gene="hemK"
    /note="PS00092 N-6 Adenine-specific DNA methylases
    signature"
    complement(1279..1288)
    /note="Core DNA uptake sequence: gccgtctgaa"
    /label=DUS
    1565..1568
    /gene="NMA0370"
    1578..2966
    /gene="NMA0370"
    1578..2966
    /gene="NMA0370"
    /note="NMA0370, probable integral membrane protein, len:
    462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
    P44640 hypothetical protein HI0325 (450 aa), fasta_scores;
    E(): 0, 49.8% identity in 464 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="putative integral membrane protein"
    /protein_id="CAB83671.1"
    /db_xref="GI:7379123"
    /db_xref="SPTREMBL:Q9JWH5"
    /translation="MNAVVAIVIMLVLSIRVHVVLSTIGAFVGGAVAGMPLONIA
    DRAGQVSQAGIIPVENKLEGGAIALSVAMLGAFAMATHSGLPQOLAGAVVRKLR
    GMPDSVRSRGVVKWLLSIIILWGMQSNIIPIHIAFIMVIFPELLVFNKLRDNR
    RLIACTVITGLVTYMFPGFAGIFNEILIGNHISAAPQLDVKVNNVNAALPDL
    MYADSLGLAFHYRKPLRYQSNNDAGNADRPPQSAVRSALAAIAVAFALQ
    GMGLDVLGAMLGAFVLMGLVINRDKANDVFGEIGKMMAMVGFIMIAQGFAAVMA
    TGHQIQLVRSMAIPGNSKGMALMVLVGLLVTMGIGSFSFTPIIAIYVPLCVGL
    GSPLATVAIVGTAGALGDAGSPASDSTIGPTMGLNADGQHDHRSVPTFIHYNIP
    LLIAGNIAMVL"
    2967..3191
    /gene="slyX"
    2967..3191
    /gene="slyX"
    /note="NMA0371, slyX, SLX protein homolog, len: 74 aa;
    similar to SLX_HAEIN P44759 SLX protein homolog (73 aa),
    fasta_scores; E(): 0.0017, 33.8% identity in 74 aa
    overlap, and SLX_ECOLI P30857 SLX protein (72 aa), fasta
    scores; E(): 0.24, 32.4% identity in 68 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="SLX protein homolog"
    /protein_id="CAB83672.1"
    /db_xref="GI:7379124"
    /db_xref="SPTREMBL:Q9JWH4"
    /translation="MDAVQEFHREITELEIQSALQEDVIAGLNAMVAELRQTLDLQQA
    QLRLYXQMDRNPDAQEYSLRDEIPHY"
    complement(3271..3423)
    /gene="NMA0372"
    complement(3271..3423)
    /gene="NMA0372"
    /note="NMA0372, unknown, questionable CDS, len: 50 aa"
    /codon_start=1
    /transl_table=11
    /product="very hypothetical protein NMA0372"

```

```
misc_feature
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="SPTREMBL:Q5JWH3"
/translation="WFGRLSPGSDGFFLAVVEYFVGICLOGKTAYFRSGGRFHRI
KGILSD"
3562..3571
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
complement(3586..4356)
/gene="thiP"
complement(3586..4356)
/gene="thiP"
/note="NMA0373, thiP, probable ThiF protein, len: 256 aa;
similar to e.g. THIF_ECOLI P30138 ThiF protein (251 aa),
fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
protein_ (249 aa), fasta scores; E(): 0, 43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopterin biosynthesis proteins). Contains
Pfam match to entry PF00899 ThiF_family, ThiF family"
/codon_start=1
/transl_table=11
/product="ThiF protein"
/protein_id="CAB83674.1"
/db_xref="GI:7379126"
/db_xref="SPTREMBL:Q5JWH2"
/translation="MTTEHNDHDAFLRYSRHLHLDIEIGGQOKLSAAHILVVGCG
GLGAALPYLAASGIGLTIADSTIVELHNLQVAFDEGVGKLKTEALADLRHIN
HTVDRTNKLDCRITGLVQAADIILDCDNYATQAVNACVQAKTPLVSGAAVR
FEGLAIVRPLDPSGVACILFDGGSASDGICSLFVFSPLVIGISTQAAEALKILL
DAGPSHRLAVRALESGWOYFDLPNPECPVCAER"
3787..3796
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
complement(3847..4257)
/gene="thiP"
/note="Pfam match to entry PF00899 ThiF_family, ThiF
family, score 186.60, E-value 4.1e-52"
complement(4418..4427)
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
4471..7224
/gene="ppc"
4471..7224
/gene="ppc"
/EC_number="4.1.1.31"
/note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
917 aa; similar to many e.g. CAPP_RHOA Q32483
phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
fasta scores; E(): 0, 43.3% identity in 928 aa overlap.
Contains 2x Pfam match to entry PF00311 PEPCase,
Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
site motif A (P-loop), and PS00393 Phosphoenolpyruvate
carboxylase active site 2"
/codon_start=1
/transl_table=11
/product="phosphoenolpyruvate carboxylase"
/protein_id="CAB83675.1"
Query Match 100.0%; Score 229; DB 1; Length 349061;
Best Local Similarity 100.0%; Pred. No. 1.7e-62;
Matches: 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTTTTCAGGTTGTCGGAAGCTCGGTACGGCAACCTGATTACGGGTGATCAGGCA 60
Db 64614 CGTTTTCAGGTTGTCGGAAGCTCGGTACGGCAACCTGATTACGGGTGATCAGGCA 64555
Qy 61 GCTTGAACATTCGACGGCAGCGGATATGTTATCCCATGAGTCGCGTAAAAAC 120
Db 64554 GCTTGAACATTCGACGGCAGCGGATATGTTATCCCATGAGTCGCGTAAAAAC 64495
Qy 121 CAATAAAGACAAATTATAGATGATGTCGGGAAGATGCCCGACCGCAAGACTATGCAAAA 180
|||||
```

```
Db 64494 CAATAAAGACAAATTATAGATGATGTCGGGAAGATGCCCGACCGCAAGACTATGCAAAA 64435
Qy 181 TATGAAAAACCAAGTACGCGGATCAGGCATGAGCATGATGATGATCAATCCG 229
|||||
Db 64434 TATGAAAAACCAAGTACGCGGATCAGGCATGAGCATGATGATGATCAATCCG 64386
|||||

RESULT 5
AE002550 15053 bp DNA linear BCT 25-MAY-2000
LOCUS Neisseria meningitidis serogroup B strain MC58 section 192 of 206
DEFINITION of the complete genome.
ACCESSION AE002550 AE002098
VERSION AE002550.2 GI:7413476
KEYWORDS
SOURCE Neisseria meningitidis MC58
ORGANISM Neisseria meningitidis MC58
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 15053)
Tetzelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Winn,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
JOURNAL
MEDLINE 20175755
PUBMED 10710307
REFERENCE
2 (bases 1 to 15053)
Tetzelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Winn,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
JOURNAL
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr. Rockville, MD 20850, USA
COMMENT On Apr 4, 2000 this sequence version replaced gi:7227263.
FEATURES
source
1..15053
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
complement(74..982)
/gene="NMB2000"
complement(74..982)
/gene="NMB2000"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42327.1"
/db_xref="GI:7227264"
/translation="MNQTAINRADVTRFIFDDMEVRLHVRLENWQHIVKQKNYPA
ARRRALLGALLSGNLKNEGTLIVQVGRGLKMLVAEASDRVTRATARWDET
ASIADESLDGLGEGVFLTLQPKDGPQGVVPLEGGIAQMLNVNMXSEOLDT
HVLVSASPAAGGLVQLPEVLDSEANEHVSILARTLIRELAGLAQHVLYLEH
ETPPVFPFEFTFECSTSRGKVSMDLLMLDGEVGVVVEGGSIEVDCDFCHSKYVF
DETVDNALFGEDVGVAKGLPRHTVQ"
1227..1982
gene
```

/gene="NMB2001"
 1227..1982
 /gene="NMB2001"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42328.1"
 /db_xref="GI:7227265"
 /translation="MDSFKPAVAVLWLFVAVRALADELTNLLSSBEQILRPAED
 EQVLPIPARARAGNADDELIGSAMGLEQVLPVNPARRAGNADDELIGNAGLN
 EQVLPIPARARAGNADDELIGNAGLEIYRGGISVSGFDCSGFMQHFIRAM
 GINLPTAEQRMGTIPVARSLEQPDQMVFFRTLGSRISHVGLYIGNRRIHAPRTG
 KNIETISHSKYKWSGYAPARVKKNDPSRLN"
 2001..2219
 /gene="NMB2002"
 2001..2219
 /gene="NMB2002"
 /note="hypothetical protein; identified by Glimmer2;
 putative"
 /transl_table=11
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAF42329.1"
 /db_xref="GI:7227266"
 /translation="MSMEMPFWVDDQCIIVSCTEKVKVMSENMAELYQTAQAFEDA
 LMGGERQLRDYLLALIEGLENPYRV"
 2364..2711
 /gene="NMB2003"
 2364..2711
 /gene="NMB2003"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42330.1"
 /db_xref="GI:7227267"
 /translation="WNIRALLIILGLATGETAVFLAGIKLPQSIQVGMGVLPALLOA
 GWKTSWQQIDALMSNLTLFVPPCVAVISYLDLIADDFWSILVSASASTLCVLV
 IGKVRHWIRGIIR"
 2708..3400
 /gene="NMB2004"
 2708..3400
 /gene="NMB2004"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42331.1"
 /db_xref="GI:7227268"
 /translation="MNSILRQPSVLLFTLAVYALAIIVTRTGNICNPVLVSTIVL
 VYKILGIDVAVYHNAAQDFDWLKPVVVLAPELVNRRKIFNQLPVIQSLAGS
 IATVGTGIVFAGLGAEREVLVSLASKSVTNPIAIEITRSIGGIPAITAATVIAIGV
 QIAGYKMLKNTVMPSSVSGMLGTASHANGIAASLRSRRMAAYAGLITFNGVLTA
 LIAPLLIPVLGF"
 3467..4687
 /gene="NMB2005"
 3467..4687
 /gene="NMB2005"
 /note="similar to GB:M65216 SP:P38434 GB:S85363 PID:150243
 PID:246444 percent identity: 98.77; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="glutamate N-acetyltransferase/amino-acid
 acetyltransferase"
 /protein_id="AAF42332.1"
 /db_xref="GI:7227269"
 /translation="WAVNLAKTAEOQLPIDIGIALYTAQGVKKPGHDTLTLVAAV
 STVCAVFTTNRFCAPVHIAKSHLFDEGVRALVINTGNANAGTGAQGRIDALAVCAA

AARQIGCKPNOVLFFSTGVILEPLPADKIIAALPKMQPAPWNEAARALMTTDTVPKAA
 SREGKVGDKHTVATGTIAKSGMWHNNMATMLGIATDAKVSQVPLQMTQEIADTFE
 NTITVDDGTSTNDSFVLIATGKNSQSEIDNIADPRYLAQLKELCSLALELAQAIADRG
 EGYKFTIVRENAKTRDEARQAAYASPLVKTAFASDPNLRLLAAIGVAGVAD
 LDTDLVEMYDDILVIAEHGGRASYTEAQGVMSKAEITVRIKLHRGQAATVYTCD
 LSHGYVSINADYRS"
 4747..6156
 /gene="NMB2006"
 4747..6156
 /gene="NMB2006"
 /note="similar to PID:1652876 percent identity: 52.04;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="chloride channel protein-related protein"
 /protein_id="AAF42333.1"
 /db_xref="GI:7227270"
 /translation="MSTFPRLARKIRQTRRSKSIAPFLFLLAGSALVALTALFFA
 HLADEFALNLAQVQYFWFVAVLPLGLFIAMLRKFAPTAGSGIPQVIALSLIP
 YGAQKTRLIRLQTLILKPLTFLGMLFGASIGREGPSVGVGAAMGAWGAKKHGLA
 FKGQENDLMAAGAGLAAAFNPLAGVIFAIEELGRGIMLRWQRIILGLVASQFI
 QVAIRGNPYPSFGNGVLEHIFLWVALSGLVCGAAGLFGRLLYEFGAAFAFRKIRG
 FIRNPILLAAALMGLLALLGTGTYGKTYGVHEAAQALHGIYEAFAGAAKWLAT
 VFSYWAGVPGGIFTPSLITGAVLGEHIAIAIDISQGANIVLICMAAFLAGAQCSEIT
 SAVVMEITGGQSLLEFWNLTIACIFASQVSRQSPRPFYHSGMRFQRVLQETAQTG
 NAPPAPQTANSKTMPSEN"
 6346..7737
 /gene="NMB2007"
 /note="This region contains a match to at least one other
 gene that is not full length, and is not the result of a
 sequencing artifact; similar to GP:1742299 percent
 identity: 79.36; identified by sequence similarity;
 putative"
 7881..9524
 /gene="NMB2008"
 7881..9524
 /gene="NMB2008"
 /note="similar to PID:862629 percent identity: 53.03;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ABC transporter, ATP-binding protein-related
 protein"
 /protein_id="AAF42334.1"
 /db_xref="GI:7227271"
 /translation="MDEIQIPKKVELQTKLENEKIVLSKGSTIIIVANGTKRLAV
 YIEQLKEXAHRISAHRAKLPNVNKKPEKAKTVLSYQWMDGIDVNRKNYRWDN
 RXYTHLLNDFWLLQVLFQAQNNIAVANNKLNKNEKVTNKTLDILOEAWETLLPH
 NKLHITADDIQVSDVNBELYSASNDGERALFVILGOVLSVDQSVLIDPEPELPH
 HXSLISNLWDKIEELPDCSFLIITHDIFPATRVAKKXVIRNYIPTPAWDISEVPEES
 NDEBITIWLRSRPIFLVEGNNSLDIATRYCIPDWIILPKGRCKDVIOVSSELK
 KUSNEMPLNLKCGIIVDUDEREIEQLNNLGIYLPVSEIENLFSUTDVAKEILK
 LNQYSEELNKLNGFKSELIKYIDNELKDXLDFEVVQVRRKIDNLYKNDLSSKI
 TETDMKSLINEISTLEQKIETWISKEIENKIEQRCIEQDLDKLLTIYDNKGLLAKSA
 CVLKGMRNKHFEFSWIMRTLKGRNKDFDAIRQKPLID"
 9568..9750
 /gene="NMB2009"
 /note="This region contains a gene with one or more
 premature stops or frameshifts, and is not the result of a
 sequencing artifact; similar to SP:Q09530 PID:733602
 percent identity: 61.70; identified by sequence
 similarity; putative"
 9962..11539

Query Match 99.3%; Score 227.4; DB 1; Length 15053;
 Best Local Similarity 99.6%; Pred. No. 3.4e-62;
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGTTTCAGGTTGTGCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATGAGGCA 60
 |||||
 DB 11433 CGGTTTCAGGTTGTGCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATGAGGCA 11492
 |||||
 QY 61 GCTTGAACATTGCGGACGGCAAGGCGGAATATGTTTATCCGCAATGATGGCGTAAAAAC 120

```

Db 11493 GCTTGAACATTCCGACGGCGAAGCGGAATATGTTATCCCAATGATGGCGTAAAAAC 11552
QY 121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 180
Db 11553 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 11612
QY 181 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCGATCCCAATCCG 229
Db 11613 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCCCAATCCG 11661

RESULT 6
AX044035 172325 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION
Sequence 114 from Patent WO0066791.
ACCESSION
AX044035
VERSION
AX044035.1 GI:11342919
KEYWORDS
Neisseria meningitidis
ORGANISM
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 114 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Location/Qualifiers
1..172325
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
227325 172325 bases"
BASE COUNT 43072 a 47583 c 41465 g 40205 t
ORIGIN
Query Match 99.3%; Score 227.4; DB 6; Length 172325;
Best Local Similarity 99.6%; Pred. No. 4.9e-62;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTCGCGAAGCTCGGTAACGGGCAACTGATTACGGGTGATCGAGCA 60
Db 24545 CGTTTCAGGTTGTCGCGAAGCTCGGTAACGGGCAACTGATTACGGGTGATCGAGCA 24604
QY 61 GCTTGAACATTCGCGACGCGGAGGATGTTATCCGCAATGATGCGGTAATAAAC 120
Db 24605 GCTTGAACATTCGCGACGCGGAGGATGTTATCCGCAATGATGCGGTAATAAAC 24664
QY 121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 180
Db 24665 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 24724
QY 181 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCCCAATCCG 229
Db 24725 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCCCAATCCG 24773

RESULT 7
AX044034 349980 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION
Sequence 113 from Patent WO0066791.
ACCESSION
AX044034
VERSION
AX044034.1 GI:11342918
KEYWORDS
Neisseria meningitidis
ORGANISM
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 114 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
227325 172325 bases"
BASE COUNT 86473 a 95646 c 85908 g 81953 t
ORIGIN
Query Match 99.3%; Score 227.4; DB 6; Length 349980;
Best Local Similarity 99.6%; Pred. No. 5.5e-62;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTCGCGAAGCTCGGTAACGGGCAACTGATTACGGGTGATCGAGCA 60
Db 324545 CGTTTCAGGTTGTCGCGAAGCTCGGTAACGGGCAACTGATTACGGGTGATCGAGCA 324604
QY 61 GCTTGAACATTCGCGACGCGGAGGATGTTATCCGCAATGATGCGGTAATAAAC 120
Db 324605 GCTTGAACATTCGCGACGCGGAGGATGTTATCCGCAATGATGCGGTAATAAAC 324664
QY 121 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 180
Db 324665 CAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGACCAAGACTATGCAAAA 324724
QY 181 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCCCAATCCG 229
Db 324725 TATGAAAAACCAAGTACCGCGATCAGGCATGGATCGATCCCAATCCG 324773

RESULT 8
AF169440/c 349 bp DNA linear BCT 09-AUG-2000
LOCUS
DEFINITION
Neisseria meningitidis strain Z2491 clone Bm013 unknown sequence.
ACCESSION
AF169440
VERSION
AF169440.1 GI:9754648
KEYWORDS
Neisseria meningitidis
ORGANISM
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 349)
Perrin, A., Nassif, X. and Tinsley, C. R.
Identification of regions of the chromosome of Neisseria
meningitidis and Neisseria gonorrhoeae which are specific to
pathogenic Neisseriae
Unpublished
2 (bases 1 to 349)
Perrin, A., Nassif, X. and Tinsley, C. R.
Direct Submission
Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vauguard, Paris 75015, France
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Location/Qualifiers
1..349
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"

```

```

/strain="Z2491"
/db_xref="taxon:487"
/clone="Em013"
BASE COUNT      76 a      97 c      85 g      91 t
ORIGIN

Query Match      88.4%; Score 202.4; DB 1; Length 349;
Best Local Similarity 99.5%; Pred. No. 2.4e-54;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 60
Db 204 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 145

QY 61 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Db 144 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 85

QY 121 CAATAAGACAAATTTAGATGATGTCGGGGAAGATGCCCGACCGCAACAGACTATGCAAAA 180
Db 84 CAATAAGACAAATTTAGATGATGTCGGGGAAGAGGCCCGACCGCAACAGACTATGCAAAA 25

QY 181 TATGAAACCAAGTACCGGATC 204
Db 24 TATGAAACCAAGTACCGGATC 1

RESULT 9
AF169418/c
LOCUS
DEFINITION Neisseria gonorrhoeae strain FA1090 clone Bg007 unknown sequence.
ACCESSION AF169418
VERSION AF169418.1 GI:9754626
KEYWORDS
SOURCE
ORGANISM Neisseria gonorrhoeae
          Neisseria gonorrhoeae
          Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
          Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 400)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Identification of regions of the chromosome of Neisseria
        meningitidis and Neisseria gonorrhoeae which are specific to
        pathogenic Neisseriae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 400)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
        rue de Vaugirard, Paris 75015, France
FEATURES
          Location/Qualifiers
            source
              1..400
              /organism="Neisseria gonorrhoeae"
              /mol_type="genomic DNA"
              /strain="FA1090"
              /db_xref="taxon:485"
              /clone="Bg007"
BASE COUNT      87 a      104 c      98 g      111 t
ORIGIN

Query Match      84.2%; Score 192.8; DB 1; Length 400;
Best Local Similarity 96.6%; Pred. No. 3.2e-51;
Matches 197; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 60
Db 204 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 145

QY 61 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Db 144 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 85

QY 121 CAATAAGACAAATTTAGATGATGTCGGGGAAGATGCCCGACCGCAACAGACTATGCAAAA 180
Db 84 CAATAAGACAAATTTAGATGATGTCGGGGAAGAGGCCCGACCGCAACAGACTATGCAAAA 25

/strain="Z2491"
/db_xref="taxon:487"
/clone="Em013"
BASE COUNT      76 a      97 c      85 g      91 t
ORIGIN

Query Match      88.4%; Score 202.4; DB 1; Length 349;
Best Local Similarity 99.5%; Pred. No. 2.4e-54;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 60
Db 204 CGTTTCAGGTTGTCGGAAGGCTCGGTAACGGGCAACCTGATTACGGGTGATCGAGCA 145

QY 61 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 120
Db 144 GCTTGAACATTCGCGACGCGGAGGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAAC 85

QY 121 CAATAAGACAAATTTAGATGATGTCGGGGAAGATGCCCGACCGCAACAGACTATGCAAAA 180
Db 84 CAATAAGACAAATTTAGATGATGTCGGGGAAGAGGCCCGACCGCAACAGACTATGCAAAA 25

QY 181 TATGAAACCAAGTACCGGATC 204
Db 24 TATGAAACCAAGTACCGGATC 1

RESULT 10
A68896/c
LOCUS
DEFINITION Sequence 67 from Patent WO9802547.
ACCESSION A68896
VERSION A68896.1 GI:4759815
KEYWORDS
SOURCE unidentified
          unidentified
          unclassified.
ORGANISM
REFERENCE 1 (bases 1 to 192)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
        Merker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
        SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
        APPLICATIONS
JOURNAL Patent: WO 9802547-A 67 22-JAN-1998;
        INST NAT SANTE RECH MED (FR)
        Other publication FR 2751000 19980116.
COMMENT
FEATURES
          Location/Qualifiers
            source
              1..192
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"
              38 a      51 c      39 g      64 t
BASE COUNT      38 a      51 c      39 g      64 t
ORIGIN

Query Match      70.9%; Score 162.4; DB 6; Length 192;
Best Local Similarity 93.8%; Pred. No. 2e-41;
Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 14 TCGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATGCGAGGCTTGAACATTGC 73
Db 192 TCTCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATGCGAGGCTTGAACATTGC 133

QY 74 CGACGGCAAGCGGGAATA-TGTTATCCGCAATGAGTGGGTHAAACCAATAAAGACAA 132
Db 132 CGACGGCAAGCGGGAATA-TGTTATCCGCAATGAGTGGGTHAAACCAATAAAGACAA 73

QY 133 ATTATAGATGATGTCGGGGAAGATGCCCGACCGCAAGACTATGCAAAATATGAAAAACCA 192
Db 72 ATTATAGATGATGTCGGGGAAGATGCCCGACCGCAAGACTATGCAAAATATGAAAAACCA 13

QY 193 AGTACGCGGATC 204
Db 12 AGTACGCGGATC 1

RESULT 11
BD062996/c
LOCUS
DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis
        species bacteria, method for obtaining them and their biological
        applications.
ACCESSION BD062996
VERSION BD062996.1 GI:22608599
KEYWORDS JP 2001504884-A/58.
SOURCE unidentified
          unidentified
          unclassified.
ORGANISM
REFERENCE 1 (bases 1 to 192)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
        Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
        species bacteria, method for obtaining them and their biological
        applications

```

JOURNAL Patent: JP 2001504684-A 58 10-APR-2001;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
 PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
 SMITHKLINE BEECHAM.
 COMMENT PN JP 2001504684-A/58
 PD 10-APR-2001
 PF 11-JUL-1997 JP 19980505685
 PR 12-JUL-1996 FR 96/08768
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
 CARLA VINALS,
 PETRA MERKER
 PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES source
 1..192
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 38 a 51 c 39 g 64 t
 BASE COUNT
 ORIGIN
 Query Match 70.9%; Score 162.4; DB 6; Length 192;
 Best Local Similarity 93.8%; Pred. No. 2e-41;
 Matches 180; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 14 TCGGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGCGAGCTTGAACATTGC 73
 Db 192 TCTCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGCGAGCTTGAACATTGC 133
 QY 74 CGACGGCAAGCGGGAATA-TGTTATTCGCAATGAGTGGCGTAAACCAATAAGACAA 132
 Db 132 CGACGGCAAGCGGGAATATGTTATTCGCAATGAGTGGCGTAAACCAATAAGACAA 73
 QY 133 ATTAGATGATGTCGGGAAGATCCCGACGACGACTATGCAAAATATGAARACCA 192
 Db 72 ATTAGATGATGTCGGGAAGATCCCGACGACGACTATGCAAAATATGAARACCA 13
 QY 193 AGTACGGGATC 204
 Db 12 AGTACGGGATC 1
 RESULT 12
 NMAJ5415
 LOCUS 409 bp DNA linear BCT 15-DEC-1998
 DEFINITION Neisseria meningitidis partial DNA sequence, clone hrtB.
 ACCESSION AJ005415
 VERSION AJ005415.1 GI:3212117
 KEYWORDS
 SOURCE Neisseria meningitidis
 ORGANISM Neisseria meningitidis
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 Neisseriaceae; Neisseria.
 REFERENCE 1
 AUTHORS Claus, H., Froesch, M. and Vogel, U.
 TITLE Identification of a hotspot for transformation of Neisseria meningitidis by shuttle mutagenesis using signature-tagged transposons
 JOURNAL Mol. Gen. Genet. 259 (4), 363-371 (1998)
 MEDLINE 99005251
 PUBMED 9730590
 REFERENCE 2 (bases 1 to 409)
 AUTHORS Vogel, U.
 TITLE Direct Submission
 JOURNAL Submitted (06-APR-1998) Vogel U., University of Wuerzburg,
 Institute for Hygiene and Microbiology, Josef-Schneider-Str. 2,
 97080 Wuerzburg, GERMANY
 COMMENT Partial sequence of meningococcal locus giving rise to high transformation frequency.
 FEATURES source
 Location/Qualifiers
 1..409

/organism="Neisseria meningitidis"
 /mol_type="genomic DNA"
 /strain="B1940"
 /isolate="German"
 /db_xref="taxon:487"
 /clone="hrtB"
 BASE COUNT 111 a 98 c 105 g 91 t 4 others
 ORIGIN
 Query Match 70.4%; Score 161.2; DB 1; Length 409;
 Best Local Similarity 93.0%; Pred. No. 5.4e-41;
 Matches 199; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
 QY 1 CGGTTTCAGGTTGTCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGCA 60
 Db 197 CGGTTTCAGGTTGTCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGCA 256
 QY 61 GCTTGAACATTCGCGACGCGAAGCGGGAATATGTTATCCGCAATGAGTGGC-TAAAAA 119
 Db 257 GCTTGAGCATTCGCGATCGCAAGCGGGAATATGTTATCCGCAATGAGTGGCTTAAAAA 316
 QY 120 CCAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAGACTATGCAAA 179
 Db 317 CCAATAAAGACAAATTTAGATGATGTCGGGGAAGATGCCGACCGCAAA-ANTATGCCAA 375
 QY 180 ATATGAAAAACCAAGTACCGGATCAGGCATGGA 213
 Db 376 ATATGAAAAACCAA-TTCNCGATNCAGGCATGGA 408
 RESULT 13
 A96274
 LOCUS 1275 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 307 from Patent WO924578.
 ACCESSION A96274
 VERSION A96274.1 GI:6780029
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1
 AUTHORS Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Maignani, V.
 TITLE Neisserial antigens
 JOURNAL Patent: WO 924578-A 307 20-MAY-1999;
 PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
 CHIRON SPA (IT); GRANDI GUIDO (IT); MASNIGNANI VEGA (IT)
 FEATURES
 source
 1..1275
 Location/Qualifiers
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 300 a 304 c 319 g 352 t
 ORIGIN
 Query Match: 46.7%; Score 107; DB 6; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 2.3e-23;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTTTCAGGTTGTCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGCA 60
 Db 1169 CGGTTTCAGGTTGTCGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGCA 1228
 QY 61 GCTTGAACATTCGCGACGCGAAGCGGGAATATGTTATCCGCAATGA 107
 Db 1229 GCTTGAACATTCGCGACGCGAAGCGGGAATATGTTATCCGCAATGA 1275
 RESULT 14
 AX024062
 LOCUS 1572 bp DNA linear PAT 15-SEP-2000
 DEFINITION Sequence 5 from Patent FR2785293.
 ACCESSION AX024062
 VERSION AX024062.1 GI:10184374


```

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTTTCAGTTGTCGCGAAGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGGCA 60
    |||||
Db 1469 CGGTTTCAGTTGTCGCGAAGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGGCA 1528
    |||||
QY 61 GGTGGAACATTCGCGACGGCAAGCGGAATATGTTTCCGCAATCA 107
    |||||
Db 1529 GGTGGAACATTCGCGACGGCAAGCGGAATATGTTTCCGCAATCA 1575
    |||||

Search completed: November 15, 2003, 03:33:55
Job time : 1016.71 secs

```

Search completed: November 15, 2003, 03:33:55
Job time : 1016.71 secs

XX
PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI Vinals C;
XX

KW	diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KX	ds.
KX	
OS	Neisseria meningitidis.
XX	
PN	WO200066791-A1.
DD	
PD	09-NOV-2000.
XX	
FF	08-MAR-2000; 200WO-US05928.
XX	
PR	30-APR-1999; 99US-0132068.
PR	08-OCT-1999; 99WO-US23573.
PR	28-FEB-2000; 200GB-0004695.
XX	
XX	(CHIR) CHIRON CORP.
PA	(GENO-) INST GENOMIC RES.
PA	
PI	Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI	Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI	Frazer CM, Grandi G;
XX	
DR	WPI; 2000-647603/62.
XX	
PT	Neisseria meningitidis B full length genome sequence and open reading
PT	frames are used to detect, treat and prevent Neisserial infections -
XX	
PS	Claim 7; Appendix A; 692pp; English.
XX	
CC	The present invention describes the full length genome of
CC	Neisseria meningitidis B (NMB). The sequences in AAP21544 and AAP21607
CC	to AAP21613 represent fragments of the NMB genomic sequence, as the
CC	sequence was too long to go in a record on its own it was split into 8
CC	sequences which overlap each other at the beginning and end of each
CC	sequence by 49980 bp (i.e. the last 49980 bp of AAP21544 is repeated at
CC	the beginning of AAP21607, the last 49980 bp of AAP21607 are repeated at
CC	the beginning of AAP21608, and so on). AAP21545 to AAP21588 encode the
CC	Neisseria proteins given in AAB58550 to AAB58593, and AAP21589 to
CC	AAP21606 represent PCR primers which are used in the exemplification of
CC	the present invention. The NMB genome and fragments from it have
CC	antibacterial activity, and can be used in vaccines and gene therapy.
CC	Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC	proteins can be used in compositions for treating or preventing infection
CC	due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC	presence of Neisserial bacteria or of antibodies raised to Neisserial
CC	bacteria. Computers, computer memory, computer storage medium or computer
CC	databases can be used in a search to identify open reading frames (ORFs)
CC	or coding sequences within the NMB genome. The DNA sequences provide
CC	further opportunities to find antigenic or immunogenic proteins which are
CC	more effective in vaccines than the outer membrane proteins currently
CC	used.
XX	
SQ	Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;
	Query Match 99.3%; Score 227.4; DB 21; Length 172325;
	Best Local Similarity 99.8%; Pred. No. Se-66;
	Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 CGGTTTCAGTGTGTCGGCAGGCGTCGGTAACGGCGAACCTGATTACGGTGATGCAGGCA 60
Dd	24545 CGGTTTCAGTGTGTCGGCAGGCGTCGGTAACGGCGAACCTGATTACGGTGATGCAGGCA 24604
QY	61 GCTTTGAACATTCGGACGCCAGGCGGAATATGTTTTATCGCAATGAGTGGCGTAAAC 120
Dd	24605 GCTTTGAACATTCGGACGCCAGGCGGAATATGTTTTATCGCAATGAGTGGCGTAAAC 24664
QY	121 CAATAAAGCAAATTTAGATGATGTCGGGGAAGATGCCCGACCGACAAGACTATGCAAAA 180
Dd	24665 CAATAAAGCAAATTTAGATGATGTCGGGGAAGATGCCCGACCGACAAGACTATGCAAAA 24724
QY	181 TATGAAAAACCAAGTAGCGCGATCAGCATGAGATGCACATCCAATTCG 229
Dd	24725 TATGAAAAACCAAGTAGCGCGATCAGCATGAGATGCACATCCAATTCG 24773

Db 324545 CCGTTTCAGGTTGTCGGCAAGGCTCGTAACGGCCAACTCATTACGGGTGATCGAGCA 324604
 QY 61 GCTTGAACATTCGGACGCGAAGCGGAATATGTTTATCGCAATGAGTGGCGTAAAAAC 120
 Db 324605 GCTTGAACATTCGGACGCGAAGCGGAATATGTTTATCGCAATGAGTGGCGTAAAAAC 324664
 QY 121 CAATAAAGACAAATTTAGATGATCTCGGGAAGATGCCCGACCGACAAGACTATGCAAAA 180
 Db 324665 CAATAAAGACAAATTTAGATGATCTCGGGAAGATGCCCGACCGACAAGACTATGCAAAA 324724
 QY 181 TATGAAAAACCAATGACCGGATCAGCATGCGATGCGATCCCAATCCG 229
 Db 324725 TATGAAAAACCAATGACCGGATCAGCATGCGATGCGATCCCAATCCG 324773

RESULT 5
 AAV03576/c
 ID AAV03576 standard; DNA; 192 BP.
 XX AC AAV03576;
 XX DT 22-OCT-1998 (first entry)
 XX DE Neisseria meningitidis DNA sequence B13.
 XX KW N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
 XX KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
 XX KW meningitis; ss.
 XX OS Neisseria meningitidis.
 XX PN WO9802547-A2.
 XX PD 22-JAN-1998.
 XX PF 11-JUL-1997; 97WO-FR01295.
 XX PR 12-JUL-1996; 96FR-0008768.
 XX PS (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX PI (SMIK) SMITHKLINE BEECHAM.
 XX PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
 XX PI Vinals C;
 XX PI WPI; 1998-110594/10.
 XX Genes present in Neisseria meningitidis but not other Neisseria
 XX PT species - and related host cells, RNA, anti-sense sequences,
 XX PT polypeptide(s) and antibodies, useful for diagnosing Neisseria
 XX PT meningitidis infection and in protective vaccines
 XX PS Example 4; Page 113; 150pp; French.
 XX CC AAV03575-606 represent sequences that are present in Neisseria
 XX CC meningitidis and N. gonorrhoeae but not in N. lactamica, except for the
 XX CC genes involved in biosynthesis of the capsule polysaccharide, fliA or C.

CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.

XX SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;
 Query Match 99.3%; Score 227.4; DB 21; Length 837096;
 Best Local Similarity 99.6%; Pred. No. 9.3e-66;
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGTTTCAGGTTGTCGGCAAGGCTCGTAACGGCCAACTCATTACGGGTGATCGAGCA 60
 Db 689316 CCGTTTCAGGTTGTCGGCAAGGCTCGTAACGGCCAACTCATTACGGGTGATCGAGCA 689375
 QY 61 GCTTGAACATTCGGACGCGAAGCGGAATATGTTTATCGCAATGAGTGGCGTAAAAAC 120
 Db 689376 GCTTGAACATTCGGACGCGAAGCGGAATATGTTTATCGCAATGAGTGGCGTAAAAAC 689435
 QY 121 CAATAAAGACAAATTTAGATGATCTCGGGAAGATGCCCGACCGACAAGACTATGCAAAA 180
 Db 689436 CAATAAAGACAAATTTAGATGATCTCGGGAAGATGCCCGACCGACAAGACTATGCAAAA 689495
 QY 181 TATGAAAAACCAATGACCGGATCAGCATGCGATGCGATCCCAATCCG 229
 Db 689496 TATGAAAAACCAATGACCGGATCAGCATGCGATGCGATCCCAATCCG 689544

RESULT 6
 AAV03576/c
 ID AAV03576 standard; DNA; 192 BP.
 XX AC AAV03576;
 XX DT 22-OCT-1998 (first entry)
 XX DE Neisseria meningitidis DNA sequence B13.
 XX KW N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
 XX KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
 XX KW meningitis; ss.
 XX OS Neisseria meningitidis.
 XX PN WO9802547-A2.
 XX PD 22-JAN-1998.
 XX PF 11-JUL-1997; 97WO-FR01295.
 XX PR 12-JUL-1996; 96FR-0008768.
 XX PS (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX PI (SMIK) SMITHKLINE BEECHAM.
 XX PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
 XX PI Vinals C;
 XX PI WPI; 1998-110594/10.
 XX Genes present in Neisseria meningitidis but not other Neisseria
 XX PT species - and related host cells, RNA, anti-sense sequences,
 XX PT polypeptide(s) and antibodies, useful for diagnosing Neisseria
 XX PT meningitidis infection and in protective vaccines
 XX PS Example 4; Page 113; 150pp; French.
 XX CC AAV03575-606 represent sequences that are present in Neisseria
 XX CC meningitidis and N. gonorrhoeae but not in N. lactamica, except for the
 XX CC genes involved in biosynthesis of the capsule polysaccharide, fliA or C.

Isolated nucleotide sequences of Neisseria meningitidis which can be
 used in the diagnosis and treatment of N. meningitidis infection and
 other Neisserial infections, for example, N.gonorrhoea -
 Claim 7; Page 629-865; 1760pp; English.

The present invention describes methods of obtaining immunogenic
 proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 represent specifically claimed Neisseria meningitidis genomic DNA
 sequences: AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 sequences, which are all used in the exemplification of the present
 invention. The nucleic acid sequences, protein sequences, and antibodies
 against them, can be used in the manufacture of a composition. The
 composition can be used as a medicament (or in the manufacture of a
 medicament) for treating, preventing or diagnosing infection due to
 Neisserial bacteria. For example, some of the identified proteins could
 be components of vaccines against Meningococcus B; against all serotypes;
 and/or against all pathogenic Neisseriae. Identification of sequences
 from the bacterium will also facilitate production of biological probes,
 particularly organism-specific probes. Attempts to make efficacious
 Meningococcus B vaccines have failed mainly due to antigen tolerance.

XX		The present invention describes methods of obtaining immunogenic
CC		proteins from Neisseria genomic sequences. AA81453 to AA82414
CC		represent specifically claimed Neisseria meningitidis genomic DNA
CC		sequences, AA81260 to AA81303 and AA825620 to AA825663 represent
CC		Neisseria DNA sequences and their corresponding proteins; AA81254 to
CC		AA81259 and AA81304 to AA81321 represent PCR primers used in the
CC		isolation of Neisseria meningitidis DNA sequences; and AA81322 to
CC		AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC		sequences, which are all used in the exemplification of the present
CC		invention. The nucleic acid sequences, protein sequences, and antibodies
CC		against them, can be used in the manufacture of a composition. The
CC		composition can be used as a medicament for in the manufacture of a
CC		medicament) for treating, preventing or diagnosing infection due to
CC		Neisserial bacteria. For example, some of the identified proteins could
CC		be components of vaccines against Meningococcus B; against all serotypes;
CC		and/or against all pathogenic Neissariae. Identification of sequences
CC		from the bacterium will also facilitate production of biological probes,
CC		particularly organism-specific probes. Attempts to make efficacious
CC		Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC		Multivalent vaccines have also been tried but none have successfully
CC		overcome antigenic variability. The provision of further, complete
CC		overcomes may provide an opportunity to identify secreted or surface
CC		exposed proteins that may be presumed targets for the immune system and
CC		which are not antigenically variable or at least more conserved than
CC		other more variable regions.
XX		
SQ		Sequence 366 BP, 84 A; 100 C; 97 G; 85 T; 0 other;
		Query Match 46.7%; Score 107; DB 21; Length 366;
		Best Local Similarity 100.0%; Pred. No. 3.4e-26;
		Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	CGCTTTTCAGGTTTCCGGGAAGGCTCGGTACCGGGCAACTGATTACGGTGATGCAGGCA 60
Dd	260	CGCTTTTCAGGTTTCCGGGAAGGCTCGGTACCGGGCAACTGATTACGGTGATGCAGGCA 319
QY	61	GCTTGAACATTCCGACGCCAAGCGGAATATGTTTTATCGCAATCA 107
Dd	320	GCTTGAACATTCCGACGCCAAGCGGAATATGTTTTATCGCAATCA 366
RESULT 8		
AAZ12104		
ID	AAZ12104	standard; DNA; 624 BP.
AC	AAZ12104;	
XX		
XX		
DT	08-OCT-1999	(first entry)
DE		Neisseria meningitidis partial ORF81 sequence.
XX		
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
KW	treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.	
OS	Neisseria meningitidis.	
XX		
Key	Location/Qualifiers	
FT	misc_feature 256..257	
FT	/tag= a	
FT	/note= 952 unspecified nucleotides are present	
FT	between these bases"	
XX		
PN	WO9924578-A2.	
PD	20-MAY-1999.	
XX		
PF	09-OCT-1998; 98WO-IB01665.	
XX		
PR	01-SEP-1998; 98GB-0019016.	
PR	06-NOV-1997; 97GB-0023516.	
PR	14-NOV-1997; 97GB-0024190.	
PR	1A-NOV-1997; 97GB-0024186.	

PR 27-NOV-1997; 97GB-0025159.
 PR 10-DEC-1997; 97GB-0026147.
 PR 14-JAN-1998; 98GB-0000759.
 XX (CHIR-) CHIRON SPA.
 PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
 DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38650.
 XX Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX Claim 9; Page 204; 524pp; English.
 CC Nucleotide sequences AA211972-Z12358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX Sequence 624 BP; 141 A; 152 C; 159 G; 170 T; 2 other;
 SQ Query Match 46.7%; Score 107; DB 20; Length 624;
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 60
 DB 518 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 577
 OY 61 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 107
 DB 578 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 624
 RESULT 9
 AA212106
 ID AA212106 standard; DNA; 1275 BP.
 AC AA212106;
 DT 08-OCT-1999 (first entry)
 DE *Neisseria meningitidis* strain A complete ORF81 sequence.
 XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW treatment; *Neisseria* infection; meningitis; septicemia; gonorrhea; ss.
 XX *Neisseria meningitidis*.
 OS *Neisseria meningitidis*.
 XX WO9924578-A2.
 PN 20-MAY-1999.
 PD 09-OCT-1998; 98WO-IB01665.
 PF 01-SEP-1998; 98GB-0019016.
 PR 06-NOV-1997; 97GB-0023516.
 PR 14-NOV-1997; 97GB-0024190.
 PR 18-NOV-1997; 97GB-0024386.
 PR 27-NOV-1997; 97GB-0025158.
 PR 10-DEC-1997; 97GB-0026147.
 PR 14-JAN-1998; 98GB-0000759.
 XX (CHIR-) CHIRON SPA.
 PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
 DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38650.
 XX Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX Claim 9; Page 204; 524pp; English.
 CC Nucleotide sequences AA211972-Z12358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX Sequence 624 BP; 141 A; 152 C; 159 G; 170 T; 2 other;
 SQ Query Match 46.7%; Score 107; DB 20; Length 624;
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 60
 DB 518 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 577
 OY 61 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 107
 DB 578 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 624

DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38652.
 XX Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX Claim 9; Page 206; 524pp; English.
 CC Nucleotide sequences AA211972-Z12358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX Sequence 1275 BP; 300 A; 304 C; 319 G; 352 T; 0 other;
 SQ Query Match 46.7%; Score 107; DB 20; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 5.5e-26;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 60
 DB 1169 CGGTTTCAGGTTGTCGCGAAGGCTCGGTACGGCAACCTGATTACGGGTGATCGAGCA 1228
 OY 61 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 107
 DB 1229 GCTTGAACATTCGCGACGCGAAGCGGCAATATGTTATCCGCAATGA 1275
 RESULT 10
 AA215298
 ID AA215298 standard; DNA; 1572 BP.
 XX AA215298;
 AC AA215298;
 DT 04-SEP-2000 (first entry)
 DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.
 XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 OS *Neisseria meningitidis*.
 FH Key Location/Qualifiers
 FT CDS 1..1572
 FT /*tag= a
 XX WO200026375-A2.
 PN 11-MAY-2000.
 PD 28-OCT-1999; 99WO-FR02643.
 PF 30-OCT-1998; 98FR-0013693.
 PR (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PI Aujame L, Bouchardon A, Renaud-Morgenie G, Rokbi B, Naesif X;
 PI Tinsley C, Perrin A;
 XX WPI; 2000-365622/31.
 DR P-PSDB; AAY93267.
 XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 PT or preventative vaccines and for diagnosis
 XX Claim 2; Page 56-59; 187pp; French.
 XX The present sequence encodes a protein that is specific for pathogenic

CC strains of *Neisseria*. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 XX
 SQ Sequence 1572 BP; 380 A; 383 C; 395 G; 414 T; 0 other;

Query Match 46.7%; Score 107; DB 21; Length 1572;
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 60
 DB 1466 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 1525

QY 61 GCTTGACATTCGCGACGGCAAGCGGGAATATGTTTATCCGAATGA 107
 DB 1526 GCTTGACATTCGCGACGGCAAGCGGGAATATGTTTATCCGAATGA 1572

RESULT 11
 AAZ12105
 ID AAZ12105 standard; DNA; 1575 BP.
 XX
 AC AAZ12105;
 XX
 DT 08-OCT-1999 (first entry)
 XX
 DE *Neisseria meningitidis* complete ORF81 sequence.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO9924578-A2.
 XX
 PD 20-MAY-1999.
 XX
 PF 09-OCT-1998; 98WB-IB01665.
 XX
 PR 01-SEP-1998; 98GB-0019016.
 PR 06-NOV-1997; 97GB-0023516.
 PR 14-NOV-1997; 97GB-0024190.
 PR 18-NOV-1997; 97GB-0024386.
 PR 27-NOV-1997; 97GB-0025158.
 PR 10-DEC-1997; 97GB-0026147.
 PR 14-JAN-1998; 98GB-0000759.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
 XX
 DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38651.
 XX
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX
 PS Claim 9; Page 205; 524pp; English.
 XX
 CC Nucleotide sequences AAZ11972-212358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX
 SQ Sequence 1575 BP; 376 A; 378 C; 405 G; 416 T; 0 other;

Query Match 46.7%; Score 107; DB 20; Length 1575;
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 60
 DB 1469 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 1528

QY 61 GCTTGACATTCGCGACGGCAAGCGGGAATATGTTTATCCGAATGA 107
 DB 1529 GCTTGACATTCGCGACGGCAAGCGGGAATATGTTTATCCGAATGA 1575

RESULT 12
 AAZ12107
 ID AAZ12107 standard; DNA; 1575 BP.
 XX
 AC AAZ12107;
 XX
 DT 08-OCT-1999 (first entry)
 XX
 DE *Neisseria gonorrhoeae* complete ORF81 sequence.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 PN WO9924578-A2.
 XX
 PD 20-MAY-1999.
 XX
 PF 09-OCT-1998; 98WB-IB01665.
 XX
 PR 01-SEP-1998; 98GB-0019016.
 PR 06-NOV-1997; 97GB-0023516.
 PR 14-NOV-1997; 97GB-0024190.
 PR 18-NOV-1997; 97GB-0024386.
 PR 27-NOV-1997; 97GB-0025158.
 PR 10-DEC-1997; 97GB-0026147.
 PR 14-JAN-1998; 98GB-0000759.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
 XX
 DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38653.
 XX
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX
 PS Claim 9; Page 208; 524pp; English.
 XX
 CC Nucleotide sequences AAZ11972-212358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.
 XX
 SQ Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;

Query Match 43.9%; Score 100.6; DB 20; Length 1575;
 Best Local Similarity 96.3%; Pred. No. 8.8e-24;
 Matches 103; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 60
 DB 1469 CGTTTCAGGTTGTGCGAAGGCTCGGTAAACGGCAACCTGATTACGGGTGATCGAGGCA 1528

QY 61 GCTTGAACTTCGCGACGCGAAGCGGGAATATGTTTATCCCAATGA 107
 DB 1529 GCTTGAACTTCGCGAAGCGGGAATATGTTTATCCCAATGA 1575

RESULT 13

AAAI5323
 ID AAAL5323 standard; DNA; 1575 BP.
 XX
 AC AAAL5323;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE DNA encoding a polypeptide of a Neisseria pathogenic strain.
 XX
 KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
 XX
 OS Neisseria gonorrhoeae.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1575
 FT /*tag= a
 XX
 PN WO200026375-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 28-OCT-1999; 99WO-FR02643.
 XX
 PR 30-OCT-1998; 98FR-0013693.
 XX
 PA (INRM) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 DR WPI; 2000-365622/31.
 DR P-PSDB; AAY93293.
 XX
 PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
 PI or preventative vaccines and for diagnosis -
 XX
 PS Claim 4; Page 137-139; 187pp; French.

XX The present sequence encodes a protein that is specific for pathogenic
 CC strains of Neisseria. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC Neisseria infections, particularly by N. meningitidis. the
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated Neisseria strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 XX
 SQ Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;
 Query Match 43.9%; Score 100.6; DB 21; Length 1575;
 Best Local Similarity 96.3%; Pred. No. 8.8e-24;
 Matches 103; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGGTTTCAGGTTGTCGCGAAGCTCGGTAACGGCAACCTGATTACGGGTGATCGAGCA 60
 DB 1469 CGGTTTCAGGTTGTCGCGAAGCTCGGTAACGAGCAACCTGATTACGGGTGATCGAGCA 1528

QY 61 GCTTGAACTTCGCGACGCGAAGCGGGAATATGTTTATCCCAATGA 107
 DB 1529 GCTTGAACTTCGCGAAGCGGGAATATGTTTATCCCAATGA 1575

RESULT 14

ABZ41628
 ID ABZ41628 standard; DNA; 1467 BP.

AC ABZ41628;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae nucleotide sequence SEQ ID 7845.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB02069.
 XX
 PR 12-FEB-2001; 2001GB-0003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Masignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR P-PSDB; ABP80658.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PI medicament for treating or preventing N. gonorrhoeae infection -
 XX
 PS Disclosure; Page 760-761; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
 CC molecules of the invention.
 XX
 SQ Sequence 1467 BP; 356 A; 358 C; 377 G; 376 T; 0 other;
 Query Match 43.3%; Score 99.2; DB 25; Length 1467;
 Best Local Similarity 97.1%; Pred. No. 2.6e-23;
 Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CGGTTTCAGGTTGTCGCGAAGCTCGGTAACGGCAACCTGATTACGGGTGATCGAGCA 60
 DB 1364 CGGTTTCAGGTTGTCGCGAAGCTCGGTAACAGGCAACCTGATTACGGGTGATCGAGCA 1423

QY 61 GCTTGAACTTCGCGACGCGAAGCGGGAATATGTTTATCCGCA 104
 DB 1424 GCTTGAACTTCGCGAAGCGGGAATATGTTTATCCGCA 1467

RESULT 15

ABZ41663
 ID ABZ41663 standard; DNA; 3228 BP.

XX
 AC ABZ41663;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae nucleotide sequence SEQ ID 7915.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB02069.

Search completed: November 15, 2003, 00:35:30
Job time : 150.797 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1527.68 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-79
Perfect score: 229
Sequence: 1 CGGTTTCAGGTTGCGCAA.....TGGATGCAGCATCAATCCG 229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vit.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.4	17.6	1201	13 BX417776	BX417776 BX417776
2	37.2	16.2	281	10 BF870084	BF870084 IL3-RT011
C 3	35.8	15.6	1474	13 BQ073621	BQ073621 AGENCOURT
C 4	35.6	15.5	707	29 BZ719643	BZ719643 PUCF032HD

5	35	15.3	1012	13 BX338876	BX338876 BX338876
6	35	15.3	1201	13 BX356738	BX356738 BX356738
C 7	34.8	15.2	841	13 BQ792412	BQ792412 EST 8867
8	34.6	15.1	1397	10 BE908439	BE908439 601503070
9	34	14.8	1005	13 BX382908	BX382908 BX382908
C 10	33.8	14.8	659	28 AQ621897	AQ621897 HS 3107 B
C 11	33.6	14.7	809	14 CB329268	CB329268 IMAGE 1d
C 12	33.6	14.7	1006	13 BX357256	BX357256 BX357256
13	33.4	14.6	514	10 BE239977	BE239977 EST104026
14	33.4	14.6	670	12 BM765283	BM765283 K-EST0046
15	33.4	14.6	693	12 B1090806	B1090806 602855332
16	33.4	14.6	743	28 AQ916226	AQ916226 rbe00063A
17	33.4	14.6	777	12 B1087286	B1087286 602850937
C 18	33.4	14.6	937	28 BH137534	BH137534 ENTNA83TF
19	33.4	14.6	1201	13 BX418945	BX418945 BX418945
20	33.2	14.5	484	12 BM873525	BM873525 laa12h08
C 21	33	14.4	710	28 BH975720	BH975720 och85a03
22	33	14.4	1201	13 BX422257	BX422257 BX422257
C 23	32.8	14.3	943	28 A2206473	A2206473 SP 0104 B
24	32.6	14.2	525	28 AZ017263	AZ017263 RPCI-23-2
25	32.6	14.2	636	10 BG623921	BG623921 602649079
C 26	32.4	14.1	429	28 A2499482	A2499482 LM0337H14
27	32.4	14.1	445	9 AA920407	AA920407 LD24006.5
28	32.4	14.1	446	9 AA820402	AA820402 LD23994.5
29	32.4	14.1	632	9 A1542901	A1542901 SD09414.5
30	32.4	14.1	682	10 BG637449	BG637449 SD15454.5
31	32.4	14.1	885	13 BQ959338	BQ959338 AGENCOURT
C 32	32.4	14.1	974	13 BX344914	BX344914 BX344914
C 33	32.4	14.1	1101	29 CMS00218	AL061925 Drosophil
C 34	32.2	14.1	241	12 B1595134	B1595134 Hc-599.07
35	32.2	14.1	244	9 AV282672	AV282672 AV282672
C 36	32.2	14.1	484	14 CB727453	CB727453 AMGNNUC.M
C 37	32.2	14.1	669	29 CC176763	CC176763 ZMMBCC030
38	32.2	14.1	670	12 BG773270	BG773270 602722033
39	32.2	14.1	999	10 BE880396	BE880396 601490876
40	32	14.0	1044	13 BX436571	BX436571 BX436571
41	32	14.0	1101	29 CMS0381N	AL108773 Drosophil
C 42	31.8	13.9	517	12 BM794419	BM794419 K-EST0075
43	31.8	13.9	507	9 AL882926	AL882926 AL882926
44	31.8	13.9	577	29 CC126610	CC126610 ND1.12C23
45	31.8	13.9	594	14 CA048887	CA048887 ssalkh000

ALIGNMENTS

RESULT 1
BX417776
LOCUS BX417776 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE012YN05
DEFINITION S-PRIME, mRNA sequence.
ACCESSION BX417776
VERSION BX417776.1 GI:30654459
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE012CG03QPl.
Location/Qualifiers
i. :1201
/organism="Homo sapiens"


```

BASE COUNT      83 a   423 c   69 g   897 t   2 others
ORIGIN

Query Match      15.6%; Score 35.8; DB 13; Length 1474;
Best Local Similarity 52.8%; Pred. No. 49;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY      82 AGCGCGAATGTTTATCGCATAGTGGGTAAACCAATTAAGCAAAATTTAGATG 141
      |||||
Db      1248 AAGCGGAAAAGACAGAGGGAAGAAACAAATTAAGAGGATTAAGAGAGAGAGAGA 1189
      |||||

QY      142 ATGTGCGGGAAGATCCCGACCGACAGACTATGCAAAATATGAAAAACCAAGTACGGG 201
      |||||
Db      1198 GGACGGGNAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1129
      |||||

QY      202 ATCAGCGATGATGACCATCCAA 225
      |||||
Db      1128 AGAAGAGAGGGAAGAGGAGAGAAA 1105
      |||||

RESULT 4
BZ719643/c
LOCUS      BZ719643      707 bp      DNA      linear      GSS 24-FEB-2003
DEFINITION PUCFO32TD ZM_0.6_1.0 KB Zea mays genomic clone ZMBT135E15,
            genomic survey sequence.
ACCESSION      BZ719643
VERSION      BZ719643.1 GI:28508337
KEYWORDS      GSS
SOURCE      Zea mays
ORGANISM      Zea mays
REFERENCE      1 (bases 1 to 707)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
            ,A., Fraser,C.M., Yuan,F., San Miguel,P., Ma,J. and Bennetzen,J.
TITLE      Maize Genomics Consortium
JOURNAL      Unpublished
COMMENT      Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES
            source
            Location/Qualifiers
            1..707
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBT135E15"
            /clone.lib="ZM_0.6_1.0 KB"
            /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
            Cor selected genomic DNA library"
BASE COUNT      212 a   175 c   92 g   228 t
ORIGIN

Query Match      15.5%; Score 35.6; DB 29; Length 707;
Best Local Similarity 51.9%; Pred. No. 56;
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY      52 ATGAGGAGCTTGAAACATTCGACGCGCAAGCGGAATATGTTATCCGAATGATGG 111
      |||||
Db      532 AGGCGAGACTTGTGGATATTGAGGAAGAAATTGAGGACGATGTTGAGTCAGATTTTATG 473
      |||||

QY      112 CGTAAACCAATTAAGACAAATTTAGATGATGTCGGGAGAGATGCCGCCCAAGAC 171
      |||||
Db      472 AAGTAGCCCAATTAATGTTGAATTTATGTATACGATTCAGATGATGACATGATTC 413
      |||||

QY      172 TATCAAAATATGAAAAACCAAGTACGGGATCA 205
      |||||

```

```

Db      412 TATCGAGTAATGAAGTTCAGACTTCTTTGATAA 379
      |||||

RESULT 5
BX338876
LOCUS      BX338876      1012 bp      mRNA      linear      EST 02-MAY-2003
DEFINITION BX338876 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1065Y17 5-PRIME, mRNA sequence.
ACCESSION      BX338876
VERSION      BX338876.1 GI:30343727
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 1012)
AUTHORS      Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4608.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0D1065AA09P1&cluster=4608.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0D1065AA09P1.
FEATURES
            source
            Location/Qualifiers
            1..1012
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0D1065Y17"
            /tissue="PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      295 a   261 c   185 g   160 t   111 others
ORIGIN

Query Match      15.3%; Score 35; DB 13; Length 1012;
Best Local Similarity 39.2%; Pred. No. 79;
Matches 74; Conservative 25; Mismatches 90; Indels 0; Gaps 0;

QY      6 TCAGGTTGTCGGGAAGGCTCGGTAAACGGGCAACCTGATTACGGGTGATCGAGGAGCTTG 65
      |||||
Db      381 TCCGCGGCTGCGGCTGTTGCCCGATGTGCGCMTGTGTCGGCGCAGMTGAGGCTGG 440
      |||||

QY      66 ACATTCGCGAGCGCAAGCGGAATATGTTTATCCGCAATAGTGGCGGTAAACCAATA 125
      |||||
Db      441 MGAATGAAAGMTGCMGMVMAGGGGAATAAGTTGAGMGMTGMGGGGGAGMAGMAAATA 500
      |||||

QY      126 AAGACAAATTTAGATGATGTCGGGAAGATGCCCGACGACAGACTATGCAAAATATGA 185
      |||||
Db      501 TCCAGCCMTTAAAGMGCGCAAGGCGCATGMTGAAGSASTATSAASWATAASMTAAAA 560
      |||||

QY      186 AAAACCAAG 194
      |||||
Db      561 ATSMTSMAG 569
      |||||

RESULT 6
BX356738
LOCUS      BX356738      1201 bp      mRNA      linear      EST 05-MAY-2003
DEFINITION BX356738 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1016Y17 5-PRIME, mRNA sequence.
ACCESSION      BX356738

```

```

VERSION      BX356738.1  GI:30366055
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      1 (bases 1 to 1201)
COMMENT      Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4608.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1016CB09QF1&cluster=4608.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Fazaday Avenue Genoscope sequence ID : CS0D1016CB09QF1.
FEATURES
Source      Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1016YD17"
/tissue type="PLACENTA" COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT  414 a 104 c 311 g 265 t 107 others
ORIGIN
Query Match      15.3%; Score 35; DB 13; Length 1201;
Best Local Similarity 50.3%; Pred. No. 79;
Matches 83; Conservative 1; Mismatches 81; Indels 0; Gaps 0;

QY  30  ACGGCGACCTGATTACGGGTGATCGGCGAGGAGCTTGAACATTCGCGACGCGCAAGCGGAA 89
DB  399  ATGTGAGAAATGAATATGGGAGAGAGTGAGGGTGGAGATGAAGATGAGAAAGGGGAA 458
QY  90  TATGTTTATCCGATGATGTCGCTAAACCAATTAAGCAATTTAGATGATGTCGGG 149
DB  459  TAAGTTGAAGAGAGATGAAGGGGAGAGAAATATGAAGAAATAAAGAGGAGAAAG 518
QY  150  GAAGATGCCACGACGAAGACTATGCAAAATATGAAACCAAG 194
DB  519  GAGATGAGTGAAGAGATGATGAAGATAAGATAAATATGATGAAR 563

RESULT 7
LOCUS      BQ792412/c
DEFINITION vitifera Grape berries Lambda Triplex2 Library Vitis
            vinifera cDNA clone CT006G02 3', mRNA sequence.
ACCESSION  BQ792412
VERSION    BQ792412.1  GI:22007378
KEYWORDS   Vitis vinifera
SOURCE     Vitis vinifera
ORGANISM   Vitis vinifera

REFERENCE
AUTHORS    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; Vitaceae; Vitis.
1 (bases 1 to 841)
Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,
Couture,C., Dedaldecamp,F., Delrot,S., Glissant,D., Grimpiet,J.,
Hamdi,S., Romieu,C. and Terrier,N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages

JOURNAL    Unpublished
COMMENT     Contact: Hamdi S.
            UMR 619 - Equipe Biologie de la Vigne
            Universite de Bordeaux I, Institut National de la Recherche
            Agronomique
            71, Avenue Edouard Bourleaux, BP 81, 33883 Villenave D'Ornon Cedex,
            France
            Tel: 00-33-(0)5-57-12-25-50
            Fax: 00-33-(0)5-57-12-25-48
            Email: s.hamdi@bordeaux.inra.fr
            Seq primer: T7.
FEATURES
Source      Location/Qualifiers
1..841
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CT006G02"
/dev stage="veraison stage"
/clone_lib="Veraison Grape berries Lambda Triplex2
Library"
/note="Organ: Fruit without seeds; Vector: Lambda Triplex2
; Site 1: SfiIA; Site 2: SfiIB; Oriented library"
BASE COUNT  157 a 225 c 116 g 343 t
ORIGIN
Query Match      15.2%; Score 34.8; DB 13; Length 841;
Best Local Similarity 54.8%; Pred. No. 89;
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY  70  TTGCGGACGCGAAGCGGCAATATGTTTATCGCAATGATGCGCTAAACCAATAAAGA 129
DB  475  TTGGCTCTGACAAAGCTAAACCCACAAAGGAAGCATTGACAAATAGGAAATAAATA 416
QY  130  CAAATTTAGATGATGTCGGGGAAGATGCCGACCGACGAGACTATGCAAAATATGAAAA 189
DB  415  GAAAGCATGGTGTGTCGAGGATGATTTTAGTGAAAAAAGAAATAAGGAAAAAGAGAGAA 356
QY  190  CCAAGT 195
DB  355  CCAAGT 350

RESULT 8
LOCUS      BE908439
DEFINITION BE908439.1  GI:10403019
            mRNA sequence.
ACCESSION  BE908439
VERSION    BE908439.1  GI:10403019
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Place: LLN9711 row: k column: 17
            High quality sequence scop: 55.
            Location/Qualifiers
            1..1397
            /organism="Homo sapiens"

```



```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3905032"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH MGC 70"
/notes="Organ: pancreas; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT      655 a      285 c      161 g      296 t
ORIGIN
Query Match      15.1%; Score 34.6; DB 10; Length 1397;
Best Local Similarity 56.6%; Pred. No. 99;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 114 TAAAAACCAATAAGACAAATTTAGATGATCGGGGAAGATGCCCGACGACAAAGCTA 173
DB 604 TATAACAAATAATATAAAATAAAGATATCTCTCAACAATCCCAACACAACTTTC 663

QY 174 TGCATAATATGAACCAAGTACGCGATCAGCATGGATGCACGATCCCAAT 226
DB 664 TTTCACCTAGCTCAACACCGCAGACATGATTTGTATACAGTATCTAAT 716

RESULT 9
BX382908
LOCUS
DEFINITION
BX382908 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
BX382908 Homo sapiens cDNA clone CS0DJ002YC08 5-PRIME, mRNA sequence.
ACCESSION
BX382908.1 GI:30455090
VERSION
BX382908.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1005)
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3223.f,
Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ002BB04QF1.
Location/Qualifiers
1..1005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ002YC08"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      380 a      175 c      244 g      195 t      11 others
ORIGIN
Query Match      14.8%; Score 34; DB 13; Length 1005;
Best Local Similarity 50.7%; Pred. No. 1.4e+02;
Matches 76; Conservative 2; Mismatches 72; Indels 0; Gaps 0;

QY 72 CGGACGGCAAGCGGGAATATGTTTATCCGCAATGAGTGGCTAAACCAATAAGACA 131

```

```

Db 414 CKGGGGRGGAAGAGAAAATTTTTCAGACCTCAAAATGATGAAAATCCATTAGATGAC 473
QY 132 AATTATAGATGATGTCGGGAAGATGCCCGACGACAAAGACTATGCAAAATATGAAAAACC 191
Db 474 AATTCTGAGGAAGAAATGAAGATGCACCAAAACAAAAGCTTTCTTAAAAAAGAGAGAA 533
QY 192 AGTAGCGCGATCAGGCATGGATGCACGAT 221
Db 534 AAGAAACAGAAACCCAGCACAGAAATTATGAT 563

RESULT 10
AQ621897/c
LOCUS
DEFINITION
AQ621897 659 bp DNA linear GSS 16-JUN-1999
HS_3107_B1_H07_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3107 Col=13 Row=P, genomic survey
sequence.
ACCESSION
AQ621897
VERSION
AQ621897.1 GI:5084289
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 659)
AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
PUBMED
10449764
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end web server: http://www.htsc.washington.edu
Plate: 3107 row: P column: 13
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 659.
FEATURES
source
1..659
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3107 Col=13 Row=P"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/notes="Organ: sperm; Vector: pBelOAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      104 a      136 c      69 g      350 t
ORIGIN
Query Match      14.8%; Score 33.8; DB 28; Length 659;
Best Local Similarity 64.9%; Pred. No. 1.6e+02;
Matches 50; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 113 GTAAACCAATAAGACAAATTTAGATGATCGGGGAAGATGCCCGACGACAAAGACT 172
Db 571 GTAAGAAAAATTAAGAAAAATTTAAGGAAGAGGATTAAGCCCAACAAAGAAAGTT 512

QY 173 ATGCAAAATATGAAAA 189
Db 511 AAGCAGACAAAGAAAAA 495

```

RESULT 11
 CB329268/c
 LOCUS CB329268 809 bp mRNA linear EST 06-MAR-2003
 DEFINITION IMAGE id 4614232 3'end NIH_MGC_92 Homo sapiens cDNA clone
 IMAGE:4614232 3', mRNA sequence.
 ACCESSION CB329268
 VERSION CB329268.1 GI:28873252
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 809)
 AUTHORS Okajima, K.
 TITLE Unpublished, Kazuki Okajima
 JOURNAL Unpublished
 COMMENT Contact: Kazuki Okajima
 Case Western Reserve University, Genetics
 2109 Adelbert Rd, Cleveland, OH 44106, USA
 Tel: 216-368-5847
 Fax: 216-368-5857
 Email: kxo23@po.cwru.edu
 Seq primer: M13rev GGAAACAGTATGACCATG
 High quality sequence stop: 500
 POLYA=Yes
 FEATURES
 source
 Location/Qualifiers
 1..809
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4614232"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 184 a 215 c 152 g 242 t 16 others
 ORIGIN
 Query Match 14.7%; Score 33.6; DB 14; Length 809;
 Best Local Similarity 59.4%; Pred. No. 1.8e+02;
 Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 53 TGCAGGAGCTTGACATTCGCGAGCGGCGGAATGTTTATCCGCAATGAGTGGC 112
 Db 158 TGCATGCAGTTGTACATTAGTCATTAAAGATGAATTATGAGTCTTAAAGATTGA 99
 QY 113 GTAAACCAATAAGACAAATTTAGATGATGTCGG 148
 Db 98 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAACATGTCGG 63
 RESULT 12
 BX357256/c
 LOCUS BX357256 1006 bp mRNA linear EST 05-MAY-2003
 DEFINITION BX357256 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CSODI023Y014 5-PRIME, mRNA sequence.
 ACCESSION BX357256
 VERSION BX357256.1 GI:30382145
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1006)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope, sequence ID : CSODI023BH07Q1.
 FEATURES
 Location/Qualifiers
 1..1006
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI023Y014"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 220 a 175 c 187 g 221 t 203 others
 ORIGIN
 Query Match 14.7%; Score 33.6; DB 13; Length 1006;
 Best Local Similarity 33.5%; Pred. No. 1.8e+02;
 Matches 57; Conservative 37; Mismatches 76; Indels 0; Gaps 0;
 QY 6 TCAGGTTTCGCGAGGCTCGTAACGGGCAACCTGATTACGGGTGATGCAGGCGCTTG 65
 Db 771 TWWSSCGGGGGGGGGGGGAAATAAAMCCCTAAWGGGKGSGGSRRTTAKAWA 712
 QY 66 AACATTTCGCGAGCGCAAGCGGAATATGTTTATCCGCAATGAGTGGCGTAAAAACCAATA 125
 Db 711 ARRTWTTRTGGGGGAGCGGCGCCCAATTSSCCSGGCGMCGGARMCAAAWAAAAA 652
 QY 126 AAGACAAATTAGATGATGTCGGGAGAGATCCCGACCGCAAGACTATG 175
 Db 651 AWAARRAATWAAGATAACGCGGGAARAVRSGSTAAAAAAMTGACCCCKG 602
 RESULT 13
 BE239977
 LOCUS BE239977 514 bp mRNA linear EST 12-JUL-2000
 DEFINITION EST0404026 MHRP- Medicago truncatula cDNA clone pMHRP-41122, mRNA
 sequence.
 ACCESSION BE239977
 VERSION BE239977.1 GI:9056160
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE 1 (bases 1 to 514)
 AUTHORS Harrison, M.J., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S.,
 Holt, I.E., Cho, J. and Fraser, C.M.
 TITLE ESTs from phosphate-starved roots of Medicago truncatula
 JOURNAL Unpublished
 COMMENT Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N265844e
 TIGR sequence name: MTHA59TK
 More information is available at
 http://chrysa.tamu.edu/medicago
 Seq primer: SKmod (CTA GAA CTA g-g GAT CC).
 FEATURES
 source
 Location/Qualifiers
 1..514
 /organism="Medicago truncatula"
 /mol_type="mRNA"

```
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pMHRP-41122"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XLOLR"
/clone_lib="MHRP-"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; At the trifoliate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this period, they were fertilized twice weekly with 1/2 Hoaglands solutions containing 200M potassium phosphate. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the UniapK vector from Strategene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."
BASE COUNT 176 a 89 c 103 g 146 t
ORIGIN

Query Match 14.6%; Score 33.4; DB 10; Length 514;
Best Local Similarity 52.5%; Pred. No. 2.1e+02;
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 65 GAACATTCGCGAGCGGATATGTTTATCGCAATGAGTGGCGTAAACCAAT 124
Db 313 GTACTACAGATGAGAGATGATACATCTGTAACCTTTGAGTGGAGGATACCTTAT 372
QY 125 AAAGACAAATTTAGATGATGTCGGGAAGATGCCGACCGACAGCACTATGCAAAATATG 184
Db 373 AAAGACACACTTTAAGGCTCTATGATCAAAATTTACTTCTACCATCATGATATGGAC 432
QY 185 AAAACCAAGTACGGGAT 203
Db 433 AAACGAGTGCACCAAT 451

RESULT 14
BM765283
LOCUS
DEFINITION
X-EST0046949 S7SNU719s1 Homo sapiens cDNA clone S7SNU719s1-3-E09
5', mRNA sequence.
BM765283
VERSION
BM765283.1 GI:19094898
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 670)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 3 row: E column: 09
High quality sequence stop: 670.
Location/Qualifiers
1. 670
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S7SNU719s1-3-E09"
/sex="M"

/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-719"
/lab_host="Top10F"
/clone_lib="S7SNU719s1"
/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(OT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in-vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the substracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."
BASE COUNT 274 a 111 c 160 g 125 t
ORIGIN

Query Match 14.6%; Score 33.4; DB 12; Length 670;
Best Local Similarity 51.7%; Pred. No. 2.1e+02;
Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 75 GACGCGAAGCGGAATATGTTTATCGCAATGAGTGGCGTAAACCAATAGACAAAT 134
Db 248 GAGGAGGAAGAGAAAATTTTCAAGACCTCAATTTGATGAAATCCATTAGTACAAT 307
QY 135 TTAGATGATGTCGGGAAGATGCCGACCGACAGACTATGCAATATGAAAAACCAAG 194
Db 308 TCTGAGGAAGAAATGGAAGATGCACCAACAAAAGCTTTCTAAAAACACAGAAAAAG 367
QY 195 TAGCGCGATCAGGCATGGATGCACGAT 221
Db 368 AAACAGAACACGACAGAAATTATGAT 394

RESULT 15
BI090806
LOCUS
DEFINITION
602855332F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4996855 5',
RNA sequence.
BI090806
VERSION
BI090806.1 GI:14509136
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 693)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
```

BASE COUNT	287 a	119 c	150 g	137 t	ORIGIN
Query Match	14.6%	Score 33.4;	DB 12;	Length 693;	
Best Local Similarity	51.7%;	Pred. No. 2.1e+02;			
Matches	76;	Conservative 0;	Mismatches 71;	Indels 0;	Gaps 0;
75	GACGGCAGGCGGAATATGTTTATCCGCAATCAGTGGCGTAAACCAATTAAGACAAAT	134			
98	GAGGAGGAGNAGAAAATTTTCAAGACCTCAATATGATGAAATCCATTAGATGACAAT	157			
135	TTAGATGATCGCGGNAGATGCCCGACGACAGACTATGCAAAATATGAAAAACCAAG	194			
158	TCTGAGGAAGAAATGGAAGATGCACCAACAAAAGCTTTCTAAAAAACAGAGAAAAG	217			
195	TACCGGATCAGGCGATGGATGCACGAT	221			
218	AAACAGAAAACGACACAGAATTATGAT	244			

Search completed: November 15, 2003, 08:02:31
Job time : 1530.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 35.6666 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-79
Perfect score: 229
Sequence: 1 CGGTTTCAGGTTGCGGAA.....TGGATGCAATCAATCCG 229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	14.4	7218	1	US-08-232-463-14
2	31.6	13.8	552	4	US-09-370-838-266
3	31.6	13.8	1583	4	US-09-370-838-224
4	31	13.5	1626	4	US-08-134-001C-1644
5	30.6	13.4	612	4	US-08-107-532A-1427
6	29.8	13.0	1923	4	US-08-620-312D-77
7	29.4	12.8	11887	4	US-08-961-527-146
8	28.4	12.4	289	3	US-08-007-005-17
9	28.4	12.4	289	3	US-08-244-796-17
10	28.4	12.4	3729	4	US-08-107-532A-1587
11	28.2	12.3	700	3	US-08-998-416-302
12	28.2	12.3	714	3	US-08-998-416-863
13	28.2	12.3	714	3	US-08-998-416-1139
14	28.2	12.3	722	3	US-08-998-416-680
15	28.2	12.3	725	3	US-08-998-416-1051
16	28.2	12.3	761	3	US-08-998-416-382
17	28.2	12.3	1512	4	US-08-328-352-1102
18	28	12.2	3713	1	US-08-100-709-1
19	28	12.2	3713	1	US-08-176-865-1
20	28	12.2	3713	1	US-08-474-038-1
21	28	12.2	3713	2	US-08-779-046-1
22	28	12.2	3713	2	US-08-881-340-1
23	27.8	12.1	3501	1	US-08-448-170-5
24	27.8	12.1	3501	3	US-08-961-803-4
25	27.6	12.1	615	4	US-08-134-001C-474
26	27.6	12.1	1664976	4	US-08-916-421B-1
27	27.4	12.0	277	3	US-08-007-005-3

C 28	27.4	12.0	277	3	US-09-244-796-3	Sequence 3, Appli
29	27	11.8	922	4	US-09-221-017B-692	Sequence 692, App
30	27	11.8	2445	4	US-09-339-159B-25	Sequence 25, Appl
31	27	11.8	3607	1	US-08-647-351B-1	Sequence 1, Appli
32	27	11.8	3607	4	US-09-326-074-1	Sequence 1, Appli
C 33	27	11.8	3804	4	US-09-620-312D-894	Sequence 894, App
C 34	27	11.8	72928	3	US-09-009-913-1	Sequence 1, Appli
35	27	11.8	1830121	4	US-09-557-884-1	Sequence 1, Appli
36	27	11.8	1830121	4	US-09-643-990A-1	Sequence 1, Appli
37	26.8	11.7	422	3	US-09-037-990B-12	Sequence 12, Appl
38	26.8	11.7	2277	1	US-08-676-967-2	Sequence 2, Appli
39	26.8	11.7	2277	1	US-08-676-974-2	Sequence 2, Appli
40	26.8	11.7	2277	2	US-09-098-487-2	Sequence 2, Appli
C 41	26.6	11.6	784	3	US-08-961-083-151	Sequence 151, App
C 42	26.6	11.6	784	4	US-09-536-784-151	Sequence 151, App
43	26.6	11.6	1289	4	US-09-247-155-138	Sequence 138, App
C 44	26.6	11.6	4643	4	US-09-453-702B-22	Sequence 22, Appl
45	26.6	11.6	6474	4	US-08-961-527-155	Sequence 155, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pizgpt-F1s
US-08-232-463-14

Query Match 14.4%; Score 33; DB 1; Length 7218;

Best Local Similarity 1.5%; Pred. No. 0.17;
Matches 3; Conservative 126; Mismatches 76; Indels 0; Gaps 0;
QY 18 GAAGCTCGGTAAGCGCAACCTGATACGGGTGATCGAGCAGCTTGAACATTCGGCAC 77
DB 1268 RRR 1209
QY 78 GCGAAGCGGGAATGTATTTCGCCAATGAGTGGCGTAAACCAATAAAGACAAATTTA 137
DB 1208 RRR 1149
QY 138 GATGATGTCGGGAAGATCGCCGACCAAGACATGCAAAATATGAAAAACAAGTAC 197
DB 1148 RRR 1089
QY 198 CGGATCAGGATGATGATCGACATC 222
DB 1088 RRR 1064

RESULT 2
US-09-370-838-266
; Sequence 266, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1998-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 266
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-370-838-266

Query Match 13.8%; Score 31.6; DB 4; Length 552;
Best Local Similarity 47.9%; Pred. No. 0.17;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 36 AACCTGATTACGGGTGATGCGAGCAGCTTGAACATTCGCGACGCGCAAGCGGGAATATGTT 95
DB 95 AAGAAGCTGGCGCGGCGAGCGAGCACCTTCTCAGTCGCGCGTTCACAGAGAA 154
QY 96 TATCGCAATGAGTGGCGTAAACCAATAAAGACAAATTTAGATGATGTCGGGAAGAT 155
DB 155 AAGCTTGGCCAGGTGGAACACAGAAATGATGCTCCTTAGAGAACCTCTTAGCAAA 214
QY 156 GCCGACCGCAAGACTATGCAAAATATGAAAAACCAAGTACGCGGATCAGGATGATG 215
DB 215 CTGNAATGACAAATATGACAGAAAAATAATGAAACAACTGAAGTGTATTGCGAG 274
QY 216 CACGATCCAA 225
DB 275 CCAATCCAA 284

RESULT 3
US-09-370-838-224
; Sequence 224, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh

APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 224
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-224

Query Match 13.8%; Score 31.6; DB 4; Length 1583;
Best Local Similarity 47.9%; Pred. No. 0.27;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 36 AACCTGATTACGGGTGATGCGAGCAGCTTGAACATTCGCGACGCGCAAGCGGGAATATGTT 95
DB 131 AAGAAGCTGGCGCGGCGAGCGAGCACCTTCTCAGTCGCGCGTTCACAGAGAA 190
QY 96 TATCGCAATGAGTGGCGTAAACCAATAAAGACAAATTTAGATGATGTCGGGAAGAT 155
DB 191 AAGCTTGGCCAGGTGGAAGACAGAAATGATGCTCCTTAGAGAACCTCTTAGCAAA 250
QY 156 GCCGACCGCAAGACTATGCAAAATATGAAAAACCAAGTACGCGGATCAGGATGATG 215
DB 251 GCTGAATGTACCAAAATATGACAGAAAAATAATGAAACAACTGAAGTGTATTGCGAG 310
QY 216 CACGATCCAA 225
DB 311 CCAATCCAA 320

RESULT 4
US-09-134-001C-1644
; Sequence 1644, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1644
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1644

Query Match 13.5%; Score 31; DB 4; Length 1626;
Best Local Similarity 52.8%; Pred. No. 0.44;
Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 63 TTGAACATTCGCGACGCGCAAGCGGGAATATGTTTATCCGCAATGATGTCGGCGTAAACCA 122
DB 279 TTTATCATTAGTCAGTACACCGTATAATCTGTGATTGACAATGAGTATTGGGATAATCA 338
QY 123 ATAAAGACAAATTTAGATGATGTCGGGAAGATGCCCGACCGACAGACTATGCAAAATA 182
DB 339 ATACCAACAAGATAAGACAAATCAACAGTAATTTTATAAAACCACTCATTTTGAATAATGA 398
QY 183 TGAATAA 189

399 AGAACAA 405

Db

RESULT 5

US-09-107-532A-1427

Sequence 1427, Application US/09107532A

Patent No. 6581275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1427:

SEQUENCE CHARACTERISTICS:

LENGTH: 612 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...612

SEQUENCE DESCRIPTION: SEQ ID NO: 1427:

US-09-107-532A-1427

Query Match 13.4%; Score 30.6; DB 4; Length 612;

Best Local Similarity 60.0%; Pred. No. 0.39;

Matches 51; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 55 CAGGCGAGCTTGACATTCGCGACGCGGAGATATGTTATCCCAATGAGTGGCGT 114

Db 330 CAAGCAAGCAATGATATTTTACGGGAGACCTTGAACGCTTTACGGAATCATTTGCCGA 389

QY 115 AAAAAACCAATAAGCAAAATTTAGA 139

Db 390 AAAAGCAAAAGAACACAAATATACA 414

RESULT 6

US-09-620-312D-77

Sequence 77, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenchua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyen

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Dmanac, Radoje T.

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784C1F2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 77

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (112) .. (1233)

US-09-620-312D-77

Query Match 13.0%; Score 29.8; DB 4; Length 1923;

Best Local Similarity 49.7%; Pred. No. 1.2; Indels 0; Gaps 0;

Matches 76; Conservative 0; Mismatches 77;

QY 60 AGCTTGAAACATTCGCGACGCGGAGATATGTTATCCGCAATGAGTGGCGTAAAAA 119

Db 33 AGACAGCACACTGCTGACTGTTTTCAGTTCTTCTGTAACAGCAGAAAGTGCACCTACTA 92

QY 120 CCAATAAGACAAATTTAGATGATGTCGGGAGAGATGCCCGACCGACAGACTATCGAAA 179

Db 93 GGAGTAGTCAGAAATTCAAATGCTGAAGAGAAAGCCATCCATGTTTCAGAGAGAGAA 152

QY 180 ATATCAAAAACCAAGTACGGGATCAGGCATGG 212

Db 153 ACATCAAAAACCAAGGAGCAGCAGCTTTGG 185

RESULT 7

US-08-961-527-146

Sequence 146, Application US/08961527

Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961.527
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 146:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11887 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-146

Query Match 12.8%; Score 29.4; DB 4; Length 11887;
 Best Local Similarity 52.0%; Pred. No. 3.8;
 Matches 66; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 88 AATATGTTATCCGCAATGAGTCGCGTAAACCAATTAAGACAAATTTAGATGATCG 147
 DB 11498 AACCTTCTAATCTGATCACTTGGGTAAGACCAATGTACTACCTTTAGATGGAACG 11557
 QY 148 GCGAAGATGCCCGCAGCAGACACTATGCAAAATATGCAAAATCAAGTACCGGATCAGG 207
 DB 11558 GGATATATGCGACCGCAGACCACTCTCCAGATAGAGATCAAAATATAGTACTTAC 11617
 QY 208 CATGAT 214
 DB 11618 CATCAAT 11624

RESULT 8
 US-09-007-005-17
 ; Sequence 17, Application US/09007005B
 ; Patent No. 6258558
 ; GENERAL INFORMATION:
 ; APPLICANT: Szostak, Jack W.
 ; APPLICANT: Roberts, Richard W.
 ; APPLICANT: Liu, Rihe
 ; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
 ; FILE OF INVENTION: FUSIONS
 ; CURRENT APPLICATION NUMBER: US/09/007.005B
 ; CURRENT FILING DATE: 1998-01-14
 ; EARLIER APPLICATION NUMBER: 60/035,963
 ; EARLIER FILING DATE: 1997-01-27
 ; EARLIER APPLICATION NUMBER: 60/064,491
 ; EARLIER FILING DATE: 1997-11-06
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 289
 ; TYPE: RNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Translation template
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(289)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-007-005-17

Query Match 12.4%; Score 28.4; DB 3; Length 289;

Best Local Similarity 15.2%; Pred. No. 1.7;
 Matches 25; Conservative 56; Mismatches 83; Indels 0; Gaps 0;
 QY 26 GGTAAACGGGCAACCTGATTACGGGTGATGCGAGCAGCTTGAACATTCGCGACGCAAGGC 85
 DB 122 SNNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNR 181
 QY 86 GGAATATGTTATCCGCAATGAGTCGCGTAAACCAATTAAGACAAATTTAGATGATGT 145
 DB 182 SNNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNR 241
 QY 146 CGGGAAGATGCCCGACCGCAGCAAGACTATGCAAAATATGAAAAA 189
 DB 242 CRURCVRURGRCRCRUAAAAA 285

RESULT 9
 US-09-244-796-17
 ; Sequence 17, Application US/09244796
 ; Patent No. 6281344
 ; GENERAL INFORMATION:
 ; APPLICANT: Szostak, Jack W.
 ; APPLICANT: Roberts, Richard W.
 ; APPLICANT: Liu, Rihe
 ; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
 ; FILE OF INVENTION: FUSIONS
 ; CURRENT APPLICATION NUMBER: US/09/244.796
 ; CURRENT FILING DATE: 1999-02-05
 ; EARLIER APPLICATION NUMBER: 60/035,963
 ; EARLIER FILING DATE: 1997-01-27
 ; EARLIER APPLICATION NUMBER: 60/064,491
 ; EARLIER FILING DATE: 1997-11-06
 ; EARLIER APPLICATION NUMBER: 09/007,005
 ; EARLIER FILING DATE: 1998-01-14
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 289
 ; TYPE: RNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Translation template
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(289)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-244-796-17

Query Match 12.4%; Score 28.4; DB 3; Length 289;
 Best Local Similarity 15.2%; Pred. No. 1.7;
 Matches 25; Conservative 56; Mismatches 83; Indels 0; Gaps 0;
 QY 26 GGTAAACGGGCAACCTGATTACGGGTGATGCGAGCAGCTTGAACATTCGCGACGCAAGGC 85
 DB 122 SNNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNR 181
 QY 86 GGAATATGTTATCCGCAATGAGTCGCGTAAACCAATTAAGACAAATTTAGATGATGT 145
 DB 182 SNNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNR 241
 QY 146 CGGGAAGATGCCCGACCGCAGCAAGACTATGCAAAATATGAAAAA 189
 DB 242 CRURCVRURGRCRCRUAAAAA 285

RESULT 10
 US-09-107-532A-1587
 ; Sequence 1587, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO


```

;
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 1587:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3729 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...3729
; SEQUENCE DESCRIPTION: SEQ ID NO: 1587:
;
; US-09-107-532A-1587
;
; Query Match 12.4%; Score 28.4; DB 4; Length 3729;
; Best Local Similarity 56.4%; Pred. No. 5.1;
; Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
;
; Qy 63 TTGAACATTCGACGCGCAGCGCGAATATGTTTATCCGAATGAGTGGGTAAACCA 122
; Db 1138 TTGCAGATTCTCAGAAAAACGAAAGATGATGTGCTGCCAGTGAAGCTTCAAAACAT 1197
;
; Qy 123 ATAAAGACAAATTTAGATGTCGGGGAAGATG 156
; Db 1198 TATCGAAAAAATTCAGAGAGTAATGGTAGATG 1231
;
; RESULT 11
; US-08-998-416-302
; Sequence 302, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; AND USES THEREOF
;
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
;
; INFORMATION FOR SEQ ID NO: 302:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1248UP
;
; US-08-998-416-302
;
; Query Match 12.3%; Score 28.2; DB 3; Length 700;
; Best Local Similarity 49.7%; Pred. No. 2.9;
; Matches 72; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; Qy 38 CTGTATTACGGGTGATGAGGCGAGTTGAACATTCCGACGCGCAAGGGGAATATGTTTA 97
; Db 392 CTTTTCACGCTGTTTCTCCAGTCTTAGCTGCGAGAAAAAGATGTATGGCGTATA 451
;
; Qy 98 TCCGAATGAGTGGCGTAAAAACCAATAAAGACAAATTTAGATGATGTGGGGAAGATGC 157
; Db 452 GCGCGTGGCCCGCGGAAAAAATAAGAAAAATAGAAAAATAGAAAAATAGAAAAAGCGTGG 511
;
; Qy 158 CCGACCGACAAGACTATGCAATA 182
; Db 512 CCGCCCGCGGCGACAGCAAGAAAA 536
;
; RESULT 12
; US-08-998-416-863
; Sequence 863, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; AND USES THEREOF
;
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
```


Search completed: November 15, 2003, 08:09:09
Job time : 39.6686 secs

D _b	837	GAGGAGGAGAGAGAAAAATTTTCAAGACCTCAAAATTGATGAAATCCATTAGATGCAAT	896
Q _y	135	TTAGATGATGTCGGGGAAGATGCCCGACCAAGACTATGCAAAATATGAAAAACCAAG	194
D _b	897	TCTGAGGAAGAAATGGAGATGCACCAAAACAAAGCTTTCTAAAAACAGAGAAAGAAG	956
Q _y	195	TACGCGGATCAGGCATGGATCAGCAT	221
D _b	957	AAACAGAAACCCAGCACAGAATTATGAT	983

RESULT 5
US-10-108-605-348
; Sequence 348, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 348
; LENGTH: 3891
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605-348

	Query Match	24.1%	Score 32.4	DB 13	Length 3891
	Best Local Similarity	56.6%	Pred. No. 5		
	Matches 60	Conservative 0	Mismatches 46	Indels 0	Gaps 0
QY	100	CGCAATGAGTGGCGTAAAAACCAATAAAGACAAATTTAGATGATGTCGGGAAGATGCCC	159		
Db	1194	CGAAATGACTGGAGGAAGCCGCTCAAGCGCATGTCTACCGGAGATGGGAAACTGCA	1253		
QY	160	GACCGCAACACTATGCAAAATATGAAAAACCAAGTACGGGGATCA	205		
Db	1254	GACAGGCAACCAAGTCAATACAGCAACACGCGAGAAATTCCTTCA	1299		

RESULT 6
US-09-738-973-266
/ Sequence 266, Application US/09738973
/ Patent No. US20020110563A1
/ GENERAL INFORMATION:
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Mohamath, Raodoh
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Scrisst, Heather
/ APPLICANT: Indirias, Carol Yoseph
/ APPLICANT: Benson, Darin R.
/ APPLICANT: Elliot, Mark
/ APPLICANT: Manniot, Jane
/ APPLICANT: Kalos, Michael D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
/ TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.475C9
/ CURRENT APPLICATION NUMBER: US/09/738.973
/ CURRENT FILING DATE: 2000-12-14

```

; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 266
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-738-973-266

Query Match      13.8%; Score 31.6; DB 10; Length 552;
Best Local Similarity 47.9%; Pred. No. 4;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy      36 AACCTGATTACGGGTGATGCAGGCAGCTTGAA CATTCGCGACGCCAAGCGCGAATATGTT 95
Db      95 AAGAAGTGGCGGCCGACGCGAGGCACCTTCCTCAGTCGCGCGCTGCAGTTCACAGAAGAA 154
Qy      96 TATCCGCAATGAGTGGCGTAAACCAATAAAGACAAATTTAGATGATGTGCGGGGAAGAT 155
Db      155 AAGCTTTGGCGAGGCTGAGAAAGACAGAAATTCGATGCTCATTAGAGAACCTTCCTTAGCAAA 214
Qy      156 GCCCGGACCGACAAGACTATCCAAAATATGAAAAACCAAGTACGCGGATCAGGCATGGATG 215
Db      215 GCTTGATGTACCAAAATATGGACAGAAAAAATATGNAACAAACTGAAGTGTATTTCGAG 274
Qy      216 CACGATCCAA 225
Db      275 CCAAAATCCAA 284

```

```

RESULT 7
US-09-854-133-266
; Sequence 266, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Monamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 056
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-266

```

Query Match	13.8%	Score 31.6;	DB 10;	Length 552;
Best Local Similarity	47.9%;	Pred. No. 4;		
Matches	91;	Conservative 0;	Mismatches 99;	Indels 0; Gaps 0;

QY	36	AACCTGATTACGGTGATGCAGGCACGCTTGAAACATTTCGCCACGCGAAGCGGGAATATGTT	95
Db	95	AGAAAGCTGGCGCCGACGCGAGCACCTTCTCAGTCGCGCCGTGCAGTTCCACAGAGAA	154
QY	96	TATCCGAATGAGTGGCGTAAAAACCAATTAAGACAAATTTAGATGATGCGGGGAAGAT	155
Db	155	AAGCTTTGGCGAGGCTGAGAAGACAGAAATTGATGCTCACTTAGAAGAACCTCCTTAGCAAA	214
QY	156	GCCCGACCGACAGACTATGCAAAATATGAAAACCAAGTACGCGGATCAGGCATGCGATG	215
Db	215	GCTGAATGTACCAAAATATGCACAGAAAAAATATGNAACAACTGAAGTGTATTGCGAG	274
QY	216	CACGATCCAA	225
Db	275	CCAAATCCAA	284

```
RESULT 8
US-10-144-649A-266
; Sequence 266, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144.649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 266
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-649A-266

Query Match 13.8%; Score 31.6; DB 14; Length 552;
Best Local Similarity 47.9%; Pred. No. 4;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 AACCTGATTACGGTGATGACGAGGAGCTTGAACATTCGCGACGCAAGGGGGAATATGTT 95
DB 95 AAGAGCTGGCGCGGCGGAGGACCTTCTCAGTCGGCGGTGCAGTTCACAGAGAA 154

QY 96 TATCCGCAATGAGTGGCGTAAACCAATAAAGACCAATTTAGATGATGTCGGGGAAGAT 155
DB 95 AAGAGCTGGCGCGGAGGACCTTCTCAGTCGGCGGTGCAGTTCACAGAGAA 154

QY 96 TATCCGCAATGAGTGGCGTAAACCAATAAAGACCAATTTAGATGATGTCGGGGAAGAT 155
DB 155 AAGCTTGCCAGGTGAGACACAGAATTGGATGCTCCTACCTAGAGAACCTCTTAGCAAA 214

QY 156 GCCCGACCGCAAGACTATGCAAAATATGAAACCAATGACCGGATCAGGCGATG 215
DB 215 GCTGAATGTACCAATATGACAGAGAAAATAATGAACAACTGAAGTGTATTGCAG 274

QY 216 CACGATCCAA 225
DB 275 CCAAAATCCAA 284

RESULT 9
US-10-044-090-43
; Sequence 43, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044.090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1990126CB1
US-10-044-090-43

Query Match 13.8%; Score 31.6; DB 13; Length 1538;
Best Local Similarity 47.9%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 AACCTGATTACGGTGATGACGAGGAGCTTGAACATTCGCGACGCAAGGGGGAATATGTT 95
DB 99 AAGAGCTGGCGCGGAGGACCTTCTCAGTCGGCGGTGCAGTTCACAGAGAA 158

RESULT 10
US-09-738-973-224
; Sequence 224, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodch
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738.973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 224
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-224

Query Match 13.8%; Score 31.6; DB 10; Length 1583;
Best Local Similarity 47.9%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 AACCTGATTACGGTGATGACGAGGAGCTTGAACATTCGCGACGCAAGGGGGAATATGTT 95
DB 131 AAGAAGCTGGCGCGGAGGACCTTCTCAGTCGGCGGTGCAGTTCACAGAGAA 190

QY 96 TATCCGCAATGAGTGGCGTAAACCAATAAAGACCAATTTAGATGATGTCGGGGAAGAT 155
DB 191 AAGCTTGCCAGGTGAGAGACAGAATTGGATGCTCCTACTAGAGAACCTCTTAGCAAA 250

QY 156 GCCCGACCGCAAGACTATGCAAAATATGAAACCAATGACCGGATCAGGCGATG 215
DB 251 GCTGAATGTACCAATATGACAGAGAAAATAATGAACAACTGAAGTGTATTGCAG 310

QY 216 CACGATCCAA 225
DB 311 CCAAAATCCAA 320

RESULT 11
US-09-854-133-224
; Sequence 224, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodch
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
```



```

; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854.133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 224
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-224

Query Match      13.8%; Score 31.6; DB 10; Length 1583;
Best Local Similarity 47.9%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 AACCTGATTACGGGTGATCGAGGACGCTTGAACTTCGCGACGGCAAGCGGAATATGTT 95
Db 131 AAGAGCTGGCGCGCGAGCGAGGACCTTCCTCAGTCGCGCGTGCAGTTCACAGAAGAA 190
QY 96 TATCCGCAATGAGTGGCTAAACCAATAAAGACAAATTTAGATGATGTCGGGGAAGAT 155
Db 191 AAGCTTGGCCAGGCTGAGAGACAGATTTGGATGCTCACTTAGAGAACTCCTTAGCAAA 250
QY 156 GCCGACCGACAGAGACTATGCAAAATATGAAAAACCAAGTACGGCGGATCGGCATGATG 215
Db 251 CTTGATGTACAAATATGGACAGAAAAAATAATGAACAACTGAAGTGTATTGCAG 310
QY 216 CACGATCCAA 225
Db 311 CCAAATCCAA 320

RESULT 12
US-10-144-649A-224
; Sequence 224, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144.649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 224
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-224

Query Match      13.8%; Score 31.6; DB 14; Length 1583;
Best Local Similarity 47.9%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 AACCTGATTACGGGTGATCGAGGACGCTTGAACTTCGCGACGGCAAGCGGAATATGTT 95
Db 131 AAGAGCTGGCGCGCGAGCGAGGACCTTCCTCAGTCGCGCGTGCAGTTCACAGAAGAA 190
QY 96 TATCCGCAATGAGTGGCTAAACCAATAAAGACAAATTTAGATGATGTCGGGGAAGAT 155
Db 191 AAGCTTGGCCAGGCTGAGAGACAGATTTGATGCTCACTTAGAGAACTCCTTAGCAAA 250
QY 156 GCCGACCGACAGACTATGCAAAATATGAAAAACCAAGTACGGCGGATCGGCATGATG 215
Db 251 CTTGATGTACAAATATGGACAGAAAAAATAATGAACAACTGAAGTGTATTGCAG 310

us-09-764-891-9126
; Sequence 9126, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: FC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9126
; LENGTH: 12100
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-9126

Query Match      13.7%; Score 31.4; DB 11; Length 12100;
Best Local Similarity 53.7%; Pred. No. 17;
Matches 65; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 104 ATGAGTGGCGTAAAAACCAATTAAGACAAATTTAGATGATGTCGGGGAAGATGCCCGACC 163
Db 4243 AAGATGAGAGAGAGATTCAAATTAAGAAAAATTAAGAAATGATGAAGGGAATGTTACCACTGA 4302
QY 164 GACAAGACTATGCAAAATATGAAAAACCAAGTACGGCGATCAGGCATGGATGCACGATCC 223
Db 4303 CCCAGAGAAAAAATAACCAACCAAGTAACTACTACGAGCACCTCTATACACACAAACT 4362
QY 224 A 224
Db 4363 A 4363

RESULT 14
US-10-091-572-702
; Sequence 702, Application US/10091572
; Publication No. US20030054373A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA118C1
; CURRENT APPLICATION NUMBER: US/10/091,572
; CURRENT FILING DATE: 2002-03-07
; Prior application number: 09/764,850
; Prior filing date: 2001-01-17
; Prior application number: 60/179,065
; Prior filing date: 2000-01-31
; Prior application number: 60/180,628
; Prior filing date: 2000-02-04
; Prior application number: 60/214,886
; Prior filing date: 2000-06-28
; Prior application number: 60/217,487
; Prior filing date: 2000-07-11
; Prior application number: 60/225,758
; Prior filing date: 2000-08-14
; Prior application number: 60/220,963
; Prior filing date: 2000-07-26
; Prior application number: 60/217,496
; Prior filing date: 2000-07-11
; Prior application number: 60/225,447
; Prior filing date: 2000-08-14
; Prior application number: 60/218,290
; Prior filing date: 2000-07-14
; Prior application number: 60/225,757
```

PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02

PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414

;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08

Query Match 13.7%; Score 31.4; DB 14; Length 12100;

Best Local Similarity 53.7%; Pred. No. 17;

Matches 65; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 104 ATGAGTGGCGGTAATAAACCACCAATTAAGACAAATTTAGATGATGTCGGGGAAGATGCCCGACC 163

Db 4243 AAGATGAGAGAGATTCACAAATAAGAAAAATTAGAAATGATGAAGGGAATGTTACCACTGA 4302

QY 164 CACAAGACATATGCAAAATATGAAAAACCAAGTACCGGATCAGGCGATGCGATGCGACGATCC 223

Db 4303 CCCACAGAAAAAATAATACACACGATTAATCTACTACGACACCTCTATACACAAACT 4362

QY 224 A 224

Db 4363 A 4363

RESULT 15

US-09-764-864-625

;; Sequence 625, Application US/09764864

;; Patent No. US20020132753A1

;; GENERAL INFORMATION:

;; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

;; FILE REFERENCE: PT223

;; CURRENT APPLICATION NUMBER: US/09/764,864

;; CURRENT FILING DATE: 2001-01-17

;; Prior application data removed - consult PALM or file wrapper

;; NUMBER OF SEQ ID NOS: 1792

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 625

;; LENGTH: 1251

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: (1220)

;; OTHER INFORMATION: n equals a,t,g, or c

;; NAME/KEY: SITE

;; LOCATION: (1249)

;; OTHER INFORMATION: n equals a,t,g, or c

US-09-764-864-625

Query Match

Best Local Similarity 13.5%; Score 31; DB 10; Length 1251;

Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 75 GAGGCAAGCGGGAATATGTTTATCCGCAATGAGTGGCGTAAACCAATAAAGACAAAT 134
Db 808 GAGGAGGAAGAAGAAAATTTTCAAGACCTCAAAATTGATGAAAATCCATTAGATGACAAAT 867
QY 135 TTAGATGATGTCGGGGAAGATGCCCGACCGACCAAGACTATGCAAAATATGAAAAACCAAG 194
Db 868 TCTGAGGAAGAAATGGAAGATGCACCAAAACAAAAGCTTTCTTAAAAACACAGAAAAAG 927
QY 195 TACCGGATCAGGCA 209
Db 928 AAACAGAAACCAGCA 942

Search completed: November 15, 2003, 08:32:02

Job time : 133.778 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 1562.62 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-83
Perfect score: 353
Sequence: 1 CGGGAATCTCAGCAGATG.....ACGGAGGACCGTTGGCTTT 353

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.on.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.gy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	353	100.0	353	6	A68912	A68912 Sequence 83
2	353	100.0	353	6	BD063012	BD063012 DNA and s
3	286.8	81.2	852	1	AF169459	AF169459 Neisseria
4	286.8	81.2	349061	1	NMA222491	AL162753 Neisseria
5	43.6	12.4	1950	10	AF071754	AF071754 Mus muscu
6	43.4	12.3	102331	1	EC0494981	AJ494981 Escherich
7	43	12.2	7170	1	AF169965	AF169965 Bacillus
8	39	11.0	139049	9	AC007314	AC007314 Homo sapi
9	38.8	11.0	328	6	AX301007	AX301007 Sequence
10	38.6	10.9	521	9	HSAP4H02	U14607 Human alpha
11	38.4	10.9	832	6	BD110729	BD110729 EST and e
12	38.4	10.9	108490	9	AC118480	AC118480 Homo sapi
13	38.4	10.9	165789	2	AL929389	AL929389 Danio rer
14	38.2	10.8	164972	2	AC101994	AC101994 Mus muscu
15	38.2	10.8	342850	1	AP003597	AP003597 Nostoc sp
16	38	10.8	233330	2	AC112548	AC112548 Rattus no
17	38	10.8	252353	2	AC094533	AC094533 Rattus no
18	37.8	10.7	187190	2	AC116412	AC116412 Mus muscu
19	37.8	10.7	234178	2	AC127449	AC127449 Rattus no
20	37.6	10.7	81304	9	HS276A23	AL449214 Homo sapi
21	37.6	10.7	129888	9	HS197B34	AL449217 Homo sapi
22	37.6	10.7	186239	9	AC112642	AC112642 Homo sapi
23	37.6	10.7	212399	9	AC084784	AC084784 Homo sapi
24	37.2	10.5	590	8	AF541915	AF541915 Arabidops
25	37.2	10.5	126386	8	AC005169	AC005169 Arabidops
26	37.2	10.5	150820	2	AC055124	AC055124 Homo sapi
27	37.2	10.5	166805	2	AC025729	AC025729 Homo sapi
28	37.2	10.5	200638	9	AL356253	AL356253 Human DNA
29	37	10.5	110000	2	AC109410	Continuation (2 of
30	37	10.5	138416	9	AC105218	AC105218 Homo sapi
31	37	10.5	175970	9	AC097488	AC097488 Homo sapi
32	37	10.5	181845	2	AC026586	AC026586 Homo sapi
33	37	10.5	190701	9	AC022274	AC022274 Homo sapi
34	37	10.5	250029	3	AE014820	AE014820 Plasmodiu
35	36.6	10.4	175352	9	AC069394	AC069394 Homo sapi
36	36.6	10.4	193709	9	AL390838	AL390838 Human DNA
37	36.6	10.4	205919	2	AC012279	AC012279 Homo sapi
38	36.6	10.4	215849	2	AL590654	AL590654 Mus muscu
39	36.4	10.3	11050	8	AY180106	AY180106 Zea mays
40	36.4	10.3	23862	2	AY211535	AY211535 Zea mays
41	36.4	10.3	148775	2	AC109832	AC109832 Oryza sat
42	36.4	10.3	174438	2	EX248081	EX248081 Danio rer
43	36.4	10.3	192100	9	AC021810	AC021810 Homo sapi
44	36.4	10.3	210523	10	AL669945	AL669945 Mouse DNA
45	36.2	10.3	408	6	AX300942	AX300942 Sequence

ALIGNMENTS

RESULT 1
A68912
LOCUS A68912 353 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 83 from Patent WO9802547.
ACCESSION A68912
VERSION A68912.1 GI:4759831
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified

REFERENCE 1 (bases 1 to 353)
AUTHORS Nassif, X., Tinsley, C., Achtman, M., Ruelle, J., Vinals, C. and Merker, P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL

Pred. No. is the number of results predicted by chance to have a

APPLICATIONS		APPLICATIIONS	
JOURNAL	Patent: WO 9802547-A 83 22-JAN-1998;	JOURNAL	Patent: WO 9802547-A 83 22-JAN-1998;
COMMENT	INST NAT SANTE RECH MED (FR)	COMMENT	INST NAT SANTE RECH MED (FR)
FEATURES	Other publication FR 2751000 19980116.	FEATURES	Other publication FR 2751000 19980116.
source	Location/Qualifiers	source	Location/Qualifiers
	1..353		1..353
	/organism="unidentified"		/organism="unidentified"
	/mol_type="genomic DNA"		/mol_type="genomic DNA"
	/db_xref="taxon:32644"		/db_xref="taxon:32644"
BASE COUNT	119 a 63 c 82 g 89 t	BASE COUNT	119 a 63 c 82 g 89 t
ORIGIN		ORIGIN	
Query Match 100.0%; Score 353; DB 6; Length 353;		Query Match 100.0%; Score 353; DB 6; Length 353;	
Best Local Similarity 100.0%; Pred. No. 4.6e-88;		Best Local Similarity 100.0%; Pred. No. 4.6e-88;	
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CGGGAATTCGAGCAGAAATGAAGAAAGCAGCGTTGATAATTTTCATAAAGTTATTGGAAG 60	QY	1 CGGGAATTCGAGCAGAAATGAAGAAAGCAGCGTTGATAATTTTCATAAAGTTATTGGAAG 60
Db	1 CGGGAATTCGAGCAGAAATGAAGAAAGCAGCGTTGATAATTTTCATAAAGTTATTGGAAG 60	Db	1 CGGGAATTCGAGCAGAAATGAAGAAAGCAGCGTTGATAATTTTCATAAAGTTATTGGAAG 60
QY	61 AAAAAGGATTTACCGTCCATTTCGGTATTTCACATACGGCTGATTACGGAATTTCCCAAA 120	QY	61 AAAAAGGATTTACCGTCCATTTCGGTATTTCACATACGGCTGATTACGGAATTTCCCAAA 120
Db	61 AAAAAGGATTTACCGTCCATTTCGGTATTTCACATACGGCTGATTACGGAATTTCCCAAA 120	Db	61 AAAAAGGATTTACCGTCCATTTCGGTATTTCACATACGGCTGATTACGGAATTTCCCAAA 120
QY	121 GCCGTAAAGATTTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCA 180	QY	121 GCCGTAAAGATTTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCA 180
Db	121 GCCGTAAAGATTTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCA 180	Db	121 GCCGTAAAGATTTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCA 180
QY	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240	QY	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
Db	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240	Db	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
QY	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240	QY	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
Db	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240	Db	181 AGTATTCGGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
QY	241 CATTATTCGAGACACCAAGACGAAACGGATTTATGCATAGCTGTGCGGGAATTAATCTG 300	QY	241 CATTATTCGAGACACCAAGACGAAACGGATTTATGCATAGCTGTGCGGGAATTAATCTG 300
Db	241 CATTATTCGAGACACCAAGACGAAACGGATTTATGCATAGCTGTGCGGGAATTAATCTG 300	Db	241 CATTATTCGAGACACCAAGACGAAACGGATTTATGCATAGCTGTGCGGGAATTAATCTG 300
QY	301 ATATCATTGAACGATTCGGCTTGATACCTAAACCGGAGAACCGTTGGCTTT 353	QY	301 ATATCATTGAACGATTCGGCTTGATACCTAAACCGGAGAACCGTTGGCTTT 353
Db	301 ATATCATTGAACGATTCGGCTTGATACCTAAACCGGAGAACCGTTGGCTTT 353	Db	301 ATATCATTGAACGATTCGGCTTGATACCTAAACCGGAGAACCGTTGGCTTT 353
RESULT 2		RESULT 3	
BD063012		AF169459	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
DNA and specific proteins or peptides of the Neisseria meningitidis		Neisseria meningitidis strain 22491 clone Cm130 unknown sequence.	
species bacteria, method for obtaining them and their biological		Neisseria meningitidis strain 22491 clone Cm130 unknown sequence.	
applications.		Neisseria meningitidis strain 22491 clone Cm130 unknown sequence.	
BD063012		AF169459	
ACCESSION		ACCESSION	
BD063012		AF169459	
VERSION		VERSION	
BD063012.1 GI:22608615		AF169459.1 GI:9754667	
KEYWORDS		KEYWORDS	
JP 2001504684-A/74.		Neisseria meningitidis	
SOURCE		SOURCE	
unidentified		Neisseria meningitidis	
ORGANISM		ORGANISM	
unclassified.		Neisseria meningitidis	
REFERENCE		REFERENCE	
1 (bases 1 to 353)		1 (bases 1 to 852)	
Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and		Perrin,A., Nassif,X. and Tinsley,C.R.	
Merker,P.		Identification of regions of the chromosome of Neisseria	
DNA and specific proteins or peptides of the Neisseria meningitidis		meningitidis and Neisseria gonorrhoeae which are specific to	
species bacteria, method for obtaining them and their biological		pathogenic Neisseriae	
applications.		Unpublished	
BD063012		JOURNAL	
BD063012		REFERENCE	
AUTHORS		Perrin,A., Nassif,X. and Tinsley,C.R.	
TITLE		Direct Submission	
JOURNAL		JOURNAL	
Patent: JP 2001504684-A 74 10-APR-2001;		Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156	
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX		rue de Vaugirard, Paris 75015, France	
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,		Location/Qualifiers	
SMITHKLINE BEECHAM		1..852	
PN JP 2001504684-A/74		/organism="Neisseria meningitidis"	
PD 10-APR-2001		/mol_type="genomic DNA"	
PF 11-JUL-1997 JP 1998050685		/strain="22491"	
PR 12-JUL-1996 FR 96/08768		/db_xref="taxon:487"	
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI		/clone="Cm130"	
CARLA VINALS,		BASE COUNT	
PI PETRA MERKER		272 a 154 c 206 g 220 t	
PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC		ORIGIN	
Strandedness: Single;		Query Match 81.2%; Score 286.8; DB 1; Length 852;	
CC Topology: Linear;		Best Local Similarity 96.6%; Pred. No. 1.6e-69;	
		Matches 336; Conservative 0; Mismatches 7; Indels 5; Gaps 4;	

QY 1 CGGGAATTCGAGCAGATCAAGAAAGCAGGCTTGATTAATTCATAAGTTATTGGAAG 60
Db 1 CGGGAATTCGAGCAGATCAAGAAAGCAGGCTTGATTAATTCATAAGTTATTGGAAG 60
QY 61 AAAAGGATTACCGTCCATTCGCGTATTCACAAATCGGCTGATTACGGAATTCGCCAAA 120
Db 61 AAAAGGATTACCGTCCATTCGCGTATTCACAAATCGGCTGATTACGGAATTCGCCAAA 120
QY 121 GCGGTAAAGATTACGTTAAATGCAACAGAGATACCAAGAAAGCTGGAACCAAGTCA 180
Db 121 GCGGTAAAGATTACGTTAAATGCAACAGAGATACCAAGAAAGCTGGAACCAAGTCA 180
QY 181 AGTATTCGGGCAACCGCTTACGCTAGCGATGTTTGGGAATGAAAATGCTTTCCTCA 239
Db 181 AGTATTCGGGCAACCGCTTACGCTAGCGATGTTTGGGAATGAAAATGCTTTCCTCA 240
QY 240 ACATATTTCAGGACACCAAGACGAAACCGATTTTATGATAGCTGTGCGGGAATTATCT 299
Db 241 AAATATTTCAGGACACCAAGACGAAACCGATTTTATGATAGCTGTGCGGGAATTATCT 299
QY 300 GATATCA--CITGAACGATTGCG--TTGATACCTATAAAACGGAGGAACC 344
Db 300 GATATCACTTGAACGATTGCGTTTGATACCTATAAAACGGAGGAACC 347
RESULT 4
NMA222491 349061 bp DNA linear BCT 02-SEP-2002
LOCUS
DEFINITION
Neisseria meningitidis serogroup A strain 22491 complete genome;
segment 2/7.
ACCESSION
AL162753 AL157959
VERSION
AL162753.2 GI:7379120
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis Z2491
Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 349061)
Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
Hollroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skelton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
JOURNAL
Nature 404 (6777), 502-506 (2000)
MEDLINE
20222556
PUBMED
10761919
REFERENCE
2 (bases 1 to 349061)
Parkhill,J.
AUTHORS
Direct Submission
TITLE
Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
JOURNAL
Notes:
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
COMMENT
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source
1. 349061
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/note="serogroup: A"
complement(24..206)
/note="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
310.39"
repeat_unit
/label=ATR
209..212
RBS
gene
220..681

CDS
/gene="NMA0368"
220..681
/gene="NMA0368"
/note="NMA0368, probable integral membrane protein, len:
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CA883669.1"
/db_xref="GI:7379121"
/db_xref="SPTREMBL:Q9JWH7"
/translation="MQEQNKSSFPPIVMLLVSVVALWIASLNSVAFVILGHGSMGLTV
LILGSIFASIDIRYCAYANYVWLAAILVILALRKVVPVHAAFWGLALVAFSVKAVYV
DEAGTSDIVRYCAGFYLVAAFAVASIGCTFAGKPKERKAASAADGTDKDV"
671..1492
/gene="hemk"
671..1492
/genes="hemk"
/notes="NMA0369, hemk, HemK protein, len: 273 aa; similar
to e.g. HEMK ECOLI P37186 HEMK protein (277 aa), fasta
scores: E(): 0, 42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
/codon_start=1
/transl_table=11
/product="HemK protein"
/protein_id="CA883670.1"
/db_xref="GI:7379122"
/db_xref="SPTREMBL:Q9JWH6"
/translation="MTFEDKWLGLSKLPKNEARMLLOYSEYTRVQLLTRGGEMPDEV
RQADRLAQRLLNGEPAVILGAREFYGRFTVNPVSLIPRTEHLVAVLARKPEN
GRVMDLGTGSAVAVTVALERPDFAFRASDISPPALETKRNAADLGARVEFAYGSMF
DTPMSEKWDIIVSNPPIENGDKHLSQDLRFEPQIALTFSDGLSIRTLLAQGAP
DRLAEGGFLLEHGFDQGAARGVLAENGFSGVETPLDLAGLDRLVTGLKYMHLK"
831..840
/gene="hemk"
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1162..1171
/genes="hemk"
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1187..1207
/gene="hemk"
/notes="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1565..1568
1578..2366
/gene="NMA0370"
1578..2366
/genes="NMA0370"
/notes="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y125_HAEIN
P44640 hypothetical protein H10325 (450 aa), fasta_scores:
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CA883671.1"
/db_xref="GI:7379123"
/db_xref="SPTREMBL:Q9JWH5"
/translation="NNAVVAVIVMLVLSLRVHVLSLTIGAFVGGAVAGPLQNI
DAAGVQSGIIPVFNKLEGKALSYAMLGAFAMATLHSGLPQOLAGAVVRKLN
GMPPDSVRSEGKWKLLSILVNMWMSQNIPIPIAFIPVLPVLLVFNRLKIDR
RLIACVTITGLVTFYPLPYGFGAIFLANEILILGNTHSAAPQDLKNNVAAAIAPAL
KGLALLAFVHYRPRLYQSNADTAGNADANRPPSAYRSLAAVAIAVCFATOL
WYEDSLVILGAMLGEAFVWMLGVINEDKANDVFEGIKOMAYGVFIWAQGAAYWNA
TGHIOPLVSSMAIFGNSKGAALMLVGLLVLTGNGIGSSFSTILFIILAIYPLCVGL
GFSPLATVAIVGTAGALDAGSPASDSTLIGPTMGLNADGQHDHSDVPTFIHNP
LLIAGWIAAWV"
2967..3191
gene


```
CDS
complement(7353..8108)
/gene="hek"
/function="involved in adhesion"
/codon_start=1
/transl_table=11
/product="adhesin/virulence factor Hek"
/protein_id="CAD4202.1"
/db_xref="GI:23954223"
/translation="MEIEMNKVFWVSWAAACVFAANAGAKCKSGFYLTGKAGASV
VSLQDFLSGDEETSKYKGGDDHDFVFGGIAAGYDFYQPSIPVRYTELEFYARKG
ADSKYVDKDSWGGYWRDDLKNEVSVNTLMLNAYDFRDSAFPTWVSAGICVAYIH
OKTGISDWDVEYSGSGSRLSRSSADNPAWSLGVGVYDVDPDIALDLSRYLDAG
DSSVYKDEMDGKYKSEVDVKSHDMLGMVNF"
complement(8311..8679)
/feature="ORF4"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD4202.1"
/db_xref="GI:23954224"
/translation="MSRQKTVQIDSPDAKFLDINTLSRLIYERKSVIENVSDVSET
VKRVHEIEKTKLSRVYETRYSTRALAGSLWSPVILGVASVAIGHNLSTWNPYE
IGKYIKRRLSVKYKAEVL"
complement(10244..10744)
/gene="prfX"
/complement(10244..10744)
/gene="prfX"
/function="putative transcriptional regulator"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="prfX protein"
/protein_id="CAD4202.1"
/db_xref="GI:23954225"
/translation="MNNIDLEKIRHOKNDPAYPREHLLMQLCIRTNKRNQDNIS
EFLGVYGINSVWVLTFLAAASHCLSPSEISOKLOFTNITRTIDFLEKAGYVR
MDSREDRRAKISITSEGMFIQRLTLAQSMYLKEIWGLTHDEQELFEVINKLLAH
LDDVSS"
complement(11058..12065)
/gene="prfG"
/complement(11058..12065)
/gene="prfG"
/function="adhesin of P-related fimbriae"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="prfG protein"
/protein_id="CAD4203.1"
/db_xref="GI:23954226"
/translation="MKWFFAPFLSLSGCNDALANQSTIFYSPNDNIYHPOLSVKV
TDIVQFVINDASATLGVACNGFTWTHGLYSEYFAMLVVPKHSVYNGVNIYLE
LQSKGSLDAEDNDNYLTKGFAMDEVNSSGRVCFDIEKRLAWSFGVGNLNLARLP
VDLPKGVTPRVKFLRGIRQNNYDIIGRYKIPSSLMKTPFNFTNLFSIKNTGGCRP
SAQSLNHCDSLNSANNHYAACTLSVCDVPTNIRFLLSNTTAYSHGQFSYGL
GHGWSIVSINGVDTGTTWRWRAQTQNTLTGSRLYGESSKIQPGVLSGATLLML
P"
complement(12109..12612)
/gene="prfF"
/complement(12109..12612)
/gene="prfF"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="prfF protein"
/protein_id="CAD4204.1"
/db_xref="GI:23954227"
/translation="WIRLSLFISSLTSVAVLADVOINRGVNIIPPTINNGQNTIW
DFGINPEHVDNSESVEYTKTISLCPYKSGSLATKYGTWNGGQNNVLATNITHEGI
ALYQKGMSTPLTLCNGSGNGRYTAGLDIARTSTFTSVPFRNGSGLINGGFRITA
SMSIYN"
complement(12687..13208)
/gene="prfE"
/complement(12687..13208)
/gene="prfE"
/function="adenine-specific methyltransferase"
/protein_id="RAF9690.1"
/db_xref="GI:96222223"
/translation="FMKLSNPDLDWLKVENKKMLGQVFTFPQGVANLWVFGNLTKP
RTILEPCFGEVFLSIQKRAYEVANDTKIIVDPVLYERVSRKPNLELYKNDFF
DFKGVYDCVIMNPPYIROELLREKRNPRFLNKSIDIITRLLPLQYPISSRSNLVYFIK

Query Match      12.3%; Score 43.4; DB 1; Length 102331;
Best Local Similarity 55.7%; Pred. No. 0.22;
Matches 83; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 21 AAAGAAAGCAGCCTTGATAATTCATAAGATTGGAAGAAAAGGATTACCGTCCAT 80
Db 44846 AAAGAAGGCCCATTTTCAGAGTTTAAAGCTACTTAAAGAGTTAGAGTATAACTATATA 44787
QY 81 TTCCGGTATTCAATACGCTGATTACGGAATTCGCCAAAGCCGTAAAGAGTTTACGTTA 140
Db 44786 TCTTTTATAGCCAACTGCTGAGAACTATGGATTCCCAAGAAAGAAAGACTCGTGCTC 44727
QY 141 ATTGCAAAAGAAATAACCAAGAAAGCT 169
Db 44726 TTAGCTAGTCGAGTAGGTAAGTTTACCCT 44698

RESULT 7
AF169965 LOCUS BCT 20-NOV-2001
DEFINITION Bacillus sp. LU11 adenine-specific methyltransferase
            (bsplU11iIMa), cytosine-specific methyltransferase (bsplU11iIMb),
            and site-specific endonuclease (bsplU11iIR) genes, complete cds.
ACCESSION AF169965
VERSION AF169965.2 GI:13163522
KEYWORDS
SOURCE Bacillus sp. LU11
ORGANISM Bacillus sp. LU11
REFERENCE 1 (bases 1 to 7170)
AUTHORS Lepikhov,K., Tchernov,A., Zheleznaia,L., Matvienko,N., Walter,J.
            and Trautner,T.A.
TITLE Characterization of the type IV restriction modification system
            BspU11iIR from Bacillus sp. LU11
JOURNAL Nucleic Acids Res. 29 (22), 4691-4698 (2001)
MEDLINE 21570545
PUBMED 11713319
REFERENCE 2 (bases 1 to 7170)
AUTHORS Lepikhov,K.A., Tchernov,A.V. and Matvienko,N.I.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1999) Group of Molecular Genetics, Institute of
            Protein Research of RAS, Prospekt Nauki, Pushchino, Moscow Region
            142292, Russia
REFERENCE 3 (bases 1 to 7170)
AUTHORS Lepikhov,K.A., Tchernov,A.V. and Matvienko,N.I.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Group of Molecular Genetics, Institute of
            Protein Research of RAS, Prospekt Nauki, Pushchino, Moscow Region
            142292, Russia
REMARK Sequence update by submitter
COMMENT On Feb 28, 2001 this sequence version replaced gi:9622222.
FEATURES
            location/Qualifiers
            1..7170
            /organism="Bacillus sp. LU11"
            /mol_type="genomic DNA"
            /strain="LU11"
            /db_xref="taxon:100666"
            /note="thermophilic"
            1553..3259
            /gene="bsplU11iIMa"
            1553..3259
            /gene="bsplU11iIMa"
            /codon_start=1
            /transl_table=11
            /product="adenine-specific methyltransferase"
            /protein_id="RAF9690.1"
            /db_xref="GI:96222223"
            /translation="FMKLSNPDLDWLKVENKKMLGQVFTFPQGVANLWVFGNLTKP
            RTILEPCFGEVFLSIQKRAYEVANDTKIIVDPVLYERVSRKPNLELYKNDFF
            DFKGVYDCVIMNPPYIROELLREKRNPRFLNKSIDIITRLLPLQYPISSRSNLVYFIK
```

BASE COUNT		2634	a	999	c	1386	g	2151	t	
ORIGIN										
Query Match		12.2%; Score 43; DB 1; Length 7170;								
Best Local Similarity		58.0%; Pred. No. 0.29;								
Matches		76;	Conservative	0;	Mismatches	55;	Indels	0;	Gaps	0;
QY	34	TTGATAATTCATAAGTTATTGGAAGAAAGGATTACCGTCCATTTCGGTATTTCACA								93
	3635	TAGACGAATTAATAATTGCCTTTAGGGAATTTGGATATGAGTTTCAGTGGAGTGTTTAA								3694
QY	94	ATACGGCTGATTACGGAATTCACCAAGCCGTAAGATTACGTTAATTCGAACAGAA								153
	3695	ATCGGCTGATTATGGAGTACCAAGAAGAAAGGAGGATATTCATAGTTGGAACAAGAG								3754
QY	154	TAACCAAGAA								164
	3755	AAGGAATAAA								3765
RESULT 8										
AC007314/c										
LOCUS		AC007314								
DEFINITION		Homo sapiens BAC clone Rpl1-74G24 from 2, complete sequence.								
ACCESSION		AC007314								
VERSION		AC007314.3								
KEYWORDS		HTG; 5523829								
SOURCE		Homo sapiens (human)								
gene		3256..4353								
CDS		/gene="bsplU1111Mb"								
		3256..4353								
		/gene="bsplU1111Mb"								
		/codon_start=1								
		/transl_table=11								
		/product="cytosine-specific methyltransferase"								
		/protein_id="AAF89681.1"								
		/db_xref="GI:9622224"								
		/translation="MKETVVSFSGGGGLDLPKNSGFNIWAIDDKDAVLTYKENL								
		GDHILGDIYKIQGLPEADVWIGGPPQCSFLVGRKSRDDEGQVLWQYLRINEI								
		RPFQFVNVVLSAKTAEGNLVLDLIIAFREIGYEVQVSLNADGVPQRRRI								
		FIVTRREGIFNFPHTNEKNGDKLWISVEEALGDLSPNTDGIKVTPLNPI								
		OLKRENNFDYVTEHSTPKLSLELDKMIIEHIVPGNYMDVPDFVPSQRIKKFETGG								
		RTTCYGLRPDKPSYTIHTFNKPNVCNTHYKPKRLITVREALRLQGFDPDNYKKS								
		TKAGKHITVGNAPVPIISQAIAKNVLSILC"								
gene		5042..7087								
CDS		/gene="bsplU1111R"								
		5042..7087								
		/gene="bsplU1111R"								
		/codon_start=1								
		/transl_table=11								
		/product="site-specific endonuclease"								
		/protein_id="AAK1351.1"								
		/db_xref="GI:13163523"								
		/translation="MQASKRSVTWKAALNYKRNPKLVENIRSVNFSLLVVPPLPVSK								
		DSEIWDTSMLDELVLQKTMGSDLAELVHSIAVSGREHEGLVPTDLELANILAIL								
		KYLGRILNENIEICDPAAGSNLLAAIRAGGTINPKQLWANDKEQLFLELLSIRLG								
		LMFPLIVSPNTSLVTKGIDCLNKDDFGDVSVMNPYVSGWKSQTKKVAARIY								
		DISGTLSTKNVGVQVEAPSLRLVTLNVLAKDMLISVVPKQYLTAKGREAEAVRNYLL								
		NDGSLNLIYIPKAGIFEDVIVLIGKSNPDKVKIKSEIPDAEINLTKFKQG								
		LNNKTNSSHSLSAYNRELNANDLHQKXGWRNLNVTNGENIDNWINSTLIPVSE								
		ELSRVHNLKRGIGNAGADLLFINSHKLWELVKDIPKMDLYPALRLVKDINNVPV								
		NSSTVDVFLAPCEKAFQTKYENILEKILDIYKVAESLAQTKQOKIKTKBELAR								
		KILNREKTRITRISATLILPRNVRYVAVFTTQSAYSTNVIEVGTGLEQKWIT								
		FSQLLSIPQLOEVMSEKQEGARKTEIGSIDLLLPKFENIDNKIVOKLINEAKTRI								
		GFLDLCNPSTITLKLWAEVLSLSKSPREILNQALFLIEEKVNERYPYLASDDE"								

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 139049)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
MEDLINE
9847074
PUBMED
2 (bases 1 to 139049)
REFERENCE
Kou,S. and Kalicki,J.
The sequence of Homo sapiens BAC clone RP11-74G24
Unpublished
3 (bases 1 to 139049)
REFERENCE
Waterston,R.H.
Direct Submission
Submitted (16-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 139049)
REFERENCE
Waterston,R.H.
Direct Submission
Submitted (17-JUL-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 139049)
REFERENCE
Waterston,R.
Direct Submission
Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 17, 1999 this sequence version replaced gi:4662678.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0074G24

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-320M2. Actual end of this
clone is at base position 139049 of RP11-74G24.
Location/Qualifiers

FEATURES

```

source
1. 139049
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-74G24"
/clone_lib="RPCI-11"
832. 893
/rpt_family="MER1_type"
repeat_region
1376. 1594
/rpt_family="Alu"
repeat_region
2219. 2524
/rpt_family="Alu"
repeat_region
2936. 3239
/rpt_family="Alu"
repeat_region
3286. 3457
/rpt_family="MER1_type"
repeat_region
3475. 4190
/rpt_family="MER21-group"
repeat_region
4228. 4548
/rpt_family="MER1_type"
repeat_region
4543. 4861
/rpt_family="Alu"
repeat_region
4862. 4883
/rpt_family="MER1_type"
repeat_region
4884. 5184
/rpt_family="Alu"
repeat_region
5185. 5564
/rpt_family="MER1_type"
repeat_region
5565. 5875
/rpt_family="Alu"
repeat_region
5876. 6517
/rpt_family="MER1_type"
repeat_region
6518. 6820
/rpt_family="Alu"
repeat_region
6821. 7393
/rpt_family="MER1_type"
repeat_region
7394. 7676
/rpt_family="Alu"
repeat_region
7677. 7697
/rpt_family="(TAAA)n"
repeat_region
7698. 7723
/rpt_family="MER1_type"
repeat_region
7757. 8064
/rpt_family="Alu"
repeat_region
8090. 8597
/rpt_family="MaLR"
repeat_region
8696. 8728
/rpt_family="L1"
repeat_region
8729. 9027
/rpt_family="Alu"
repeat_region
9032. 9079
/rpt_family="(CA)n"
repeat_region
9080. 9206
/rpt_family="L1"
repeat_region
9210. 9352
/rpt_family="Alu"
repeat_region
9353. 9666
/rpt_family="Alu"
repeat_region
9667. 9878
/rpt_family="L1"
repeat_region
9879. 10047
/rpt_family="Alu"
repeat_region
10048. 10126
/rpt_family="(TTTC)n"
repeat_region
10127. 10271
/rpt_family="Alu"
repeat_region
10272. 10299
/rpt_family="L1"
repeat_region
10306. 10595
/rpt_family="Alu"
repeat_region

repeat_region
13524. 13656
/rpt_family="Retroviral"
repeat_region
15083. 15207
/rpt_family="MaLR"
repeat_region
15208. 15496
/rpt_family="MER4-group"
repeat_region
15497. 16118
/rpt_family="Alu"
repeat_region
16185. 16223
/rpt_family="MER4-group"
repeat_region
16437. 16580
/rpt_family="(TC)n"
repeat_region
17545. 17841
/rpt_family="MaLR"
repeat_region
17842. 18143
/rpt_family="Alu"
repeat_region
18399. 18447
/rpt_family="L1"
repeat_region
18499. 18766
/rpt_family="L1"
repeat_region
18769. 19074
/rpt_family="Alu"
repeat_region
20185. 20483
/rpt_family="L1"
repeat_region
20485. 20523
/rpt_family="L2"
repeat_region
20925. 21173
/rpt_family="MaLR"
repeat_region
21185. 21340
/rpt_family="L2"
repeat_region
21341. 21644
/rpt_family="Alu"
repeat_region
21645. 21667
/rpt_family="L2"
repeat_region
22413. 22475
/rpt_family="L2"
repeat_region
23183. 23295
/rpt_family="MIR"
repeat_region
24065. 24140
/rpt_family="MIR"
repeat_region
24589. 24734
/rpt_family="MER1_type"
repeat_region
24737. 24845
/rpt_family="L2"
repeat_region
26770. 27073
/rpt_family="Alu"
repeat_region
27395. 27631

Query Match 11.0%; Score 39; DB 9; Length 139049;
Best Local Similarity 48.8%; Pred No. 3.7;
Matches 105; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 14 CAGATGAAGAAACACGCTTGATTAATTCATAAAGTTATTGGAAGAAAAGGATTTAC 73
Db 47536 CAGTTTGAGAGATACACGCTTTATTGTGTGTCAGATCAGCAATGGAAGATTTAG 47477
QY 74 CGTCCATTTCGGTATTCACATACGCTGATTACGGATTCCTCCCAAGCCGTAAGATT 133
Db 47476 CGTCATTGTATAAGTAAGATTAGGCTAAGATGGAAGCTTTGGGAAAACAATATTT 47417
QY 134 TACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCAGTCAAGTATTCGGCAA 193
Db 47416 TACTTGAATCTGTGATAAAATCCATACCATACACCTCTCTAGAGAAAATACAAATGCT 47357
QY 194 ACGGCTTACGGTAGCCGATGTTTGGGAATGAAA 228
Db 47356 GTGATGTTATTACATAGAAATTTTAGAGAAATTCACA 47322

RESULT 9
AX301007/c AX301007 328 bp DNA linear PAT 30-NOV-2001
LOCUS

```

```

DEFINITION Sequence 78 from Patent WO0184903.
ACCESSION AX301007
VERSION AX301007.1 GI:17382272
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Donne-Gousse,C., Laudet,V. and Hanni,C.
TITLE Method for detecting and identifying the presence of biological
JOURNAL substances derived from birds, and oligonucleotides therefor
Patent: WO 0184903-A 78 15-NOV-2001;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
source
Location/Qualifiers
1..328
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="fragment d'ADN amplifie partir du genome d'oiseaux"
BASE COUNT 34 a 22 c 14 g 21 t 237 others
ORIGIN
Query Match 11.0%; Score 38.8; DB 6; Length 328;
Best Local Similarity 11.9%; Pred. No. 4.7;
Matches 36; Conservative 149; Mismatches 117; Indels 0; Gaps 0;
QY 7 TTCTGAGCAGATGAAGAAAGCAGCGCTGATATTCATTAAGTATTGGAAGAAAAG 66
Db 328 DNDKKRWSWAKGRIYRDDDBKXRYRRYDRDHYDRBNVDKTKNRNRDB 269
QY 67 GATTTACCGTCATTCGGTATTCACATACGGCTGATTACGGAATCCCAAGCCGTA 126
Db 268 DRRDTRTRKDDWDYDYNRDKTDNDWBHBBBYTRGTRCTGGGTTCTTAAATAADKW 209
QY 127 AAAGATTACGTTAATTCGAACAGATAACCAAGAAAGCTGGAACCACTCAAGTATT 186
Db 208 TRDDBRHRDHRTKHDWYNKWDHGHVNRVNDVBNRDRSRKRWTDNRN 149
QY 187 CGGCAAAAGCGGTACCGTAGCGGATGTTTGGGAATGGAATGCCATCCCAACATAT 246
Db 148 DGVRETDDYRNDNRDNDNDDBDDKDDYDGTGTHKRTGTTKHTYTKNDKDTATG 89
QY 247 TGCAGGACCAAGCAAGAAAGGATTATTGATAGCTAGCTGCGGGAATATCTCATATCA 306
Db 88 TCCDCRAGCATTCACYNRRYARCAYBWDVNVNTRGRDBDGYGGANRRRCWYACYRA 29
QY 307 CT 308
Db 28 AT 27
RESULT 10
HSAP4H02/c
LOCUS HSAP4H02 521 bp DNA linear PRI 07-JAN-1995
DEFINITION Human alpha-subunit of prollyl 4-hydroxylase gene, exon 2.
ACCESSION U14607
VERSION U14607.1 GI:602659
KEYWORDS
SEGMENT
SOURCE 2 of 16
Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS HeLaakoski,T., Veijola,J., Vuori,K., Rehn,M., Chow,L.T.,
TITLE Tailon-Miller,P., Kivirikko,K.I. and Philajaniemi,T.
Structure and expression of the human gene for the alpha subunit of
prolyl 4-hydroxylase. The two alternatively spliced types of mRNA
correspond to two homologous exons the sequences of which are
expressed in a variety of tissues
J. Biol. Chem. 269 (45), 27847-27854 (1994)
95050550
PUBMED 7961714

```

```

REFERENCE 2 (bases 1 to 521)
AUTHORS Rehn,M.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-1994) Marko Rehn, Department of Medical
Biochemistry, University of Oulu, Kajaanintie 52 A, Oulu, Finland,
Fin-90220
FEATURES
source
Location/Qualifiers
1..521
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..208
/number=1
/cons_splices=(5'site:NO, 3'site:YES)
/evidence=experimental
209..316
/product="alpha-subunit of prollyl 4-hydroxylase"
/number=2
/evidence=experimental
317..521
/number=2
/cons_splices=(5'site:YES, 3'site:NO)
/evidence=experimental
BASE COUNT 161 a 76 c 71 g 213 t
ORIGIN
Query Match 10.9%; Score 38.6; DB 9; Length 521;
Best Local Similarity 46.5%; Pred. No. 5.3;
Matches 125; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 7 TTCTGAGCAGATGAAGAAAGCAGCGCTGATATTCATTAAGTATTGGAAGAAAAG 66
Db 490 TGTATATAAAACGTAACGAAATGTCAAATATTATAAAAGGACCTTTAGGGCAGG 431
QY 67 GATTTACCGTCCATTTCCGTTATTCACATACGGCTGATTACGGAATCCCAAGCCGTA 126
Db 430 GAATAAACCTATCTACTCGAATACACTGATTGTTGTTGCTGATTTAAACACTAAA 371
QY 127 AAAGATTACGTTAATTCGAACAGATAACCAAGAAAGCTGGAACCAAGTCAAGTATT 186
Db 370 AATGGAATTAATAAATGCAAAACCAACCAACCAACCAAGTAGTAGTACCAATTG 311
QY 187 CGGCAAAAGCGGTACCGTAGCGGATGTTTGGGAATGGAATGCTTTCCCAACATTAT 246
Db 310 AAGTAAAAAGCCCTGGATGAGCCAAAGACTGGGAAGCAGAAATTCCTATAATTATAT 251
QY 247 TGCAGGACCAAGACGAAACGGATTTA 275
Db 250 ACCAGATCATCTTGGGAATTTAATTTA 222
RESULT 11
BD110729/c
LOCUS BD110729 832 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD110729
VERSION BD110729.1 GI:23205547
KEYWORDS JP 2002010789-A/2806.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 2806 15-JAN-2002;
GENSET CORP
COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/2806
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI

```



```

/rpt_family="ERV1"
1108_..2507
/rpt_family="L1"
21892_..22316
/rpt_family="L1"
22317_..22621
/rpt_family="Alu"
22622_..23339
/rpt_family="L1"
23357_..24625
/rpt_family="ERV1"
24634_..24814
/rpt_family="L1"
24854_..25176
/rpt_family="ERV1"
25275_..25822
/rpt_family="L1"
25625_..25649
/rpt_family="(TA)n"
25783_..25898
/rpt_family="L1"
25893_..25986
/rpt_family="ERV1"
25991_..26402
/rpt_family="L1"

Query Match 10.9%; Score 38.4; DB 9; Length 108490;
Best Local Similarity 51.1%; Pred. No. 5.4;
Matches 90; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 10 TGAGCAGAAATGAAGAAAGCAGCGTTGATATTTTATAAGTTATTGGAAGAAAAAGGAT 69
Db \15049 TGAGGAGAACAAAGAAATAGATTAATCAGACATAAATGAAATATAAATATAAATAGAT 15108.
QY 70 TTACCGTCATTTCGGTATTCACATACGGCTGATTACGGAATTCCTCCAAAGCGTAAAA 129
Db 15109 TAAAAAATGATAGAGCATCAACAAATCAAAAGTTGGTTTATTGAAGATAGTAATAA 15168
QY 130 GATTTCAGTTAATTCACACAGAAATACCAAGAAAAAGCTGGAACACAGTCAAGTAT 185
Db 15169 TTGTTAATCTCTAGCAAGCCAGTCAGAGAAAAAGGACATGAACCATTAATCAAT 15224

RESULT 13
AL929389/c
LOCUS
DEFINITION
Danio rerio clone CH211-51H4, *** SEQUENCING IN PROGRESS ***, 3
unordered pieces.
ACCESSION
AL929389
VERSION
AL929389.4 GI:30025397
KEYWORDS
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 165789)
AUTHORS
Andrew R.
DIRECT SUBMISSION
Submitted (22-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 19, 2003 this sequence version replaced gi:28268589.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project information
Center project name: ZC51H4
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 165008 bases at least Q40

```

Consensus quality: 165145 bases at least Q30
 Consensus quality: 165247 bases at least Q20
 Insert size: 165589; sum-of-contigs
 Insert size: 160385; 3.5% error; agarose-fp
 Quality coverage: 8.68x in Q20 bases; sum-of-contigs
 coverage: 8.96x in Q20 bases; agarose-fp

***** NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 69969: contig of 69969 bp in length
 69970 70069: gap of 100 bp
 70070 92004: contig of 21935 bp in length
 92005 92104: gap of 100 bp
 92105 165789: contig of 73685 bp in length.

Location/Qualifiers
 1.165789
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-51H4"
 /clone_lib="CHORI-211"

misc_feature
 1.69969
 /note="assembly fragment:01530
 fragment chain:1"

misc_feature
 70070..92004
 /note="assembly fragment:00845
 fragment chain:1"

misc_feature
 92105..165789
 /note="assembly fragment:00602"
 fragment chain:1

BASE COUNT 50716 a 30908 c 30918 g 53047 t 200 others

ORIGIN

Query Match 10.9%; Score 38.4; DB 2; Length 165789;
 Best local Similarity 51.1%; Pred. No. 5.4;
 Matches 90; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 CGGAATTCGAGCAGATGAAGAAAGCAGGCTGTGATTAATTCATAAATTTTGAAG 60
 Db 134880 CGACAAGTTTCAGTAATAAATAAAAAAACTGTAAATTAACAAAAAACAATAACAA 134821

QY 61 AAAAGAGTTTACCGTCCATTCGCTATTCACATACGGCTGATTACGGGAATTCGCCA 120
 Db 134820 AAAAACAATCTACTTTCGCTAAGATTTTAATTCGAATTAATGATTCCTCAGTATATAAT 134761

QY 121 GCCGTAAAGATTTACGTTTATTCGATTAATTCACACAGATAAACAAGAAAGCTGGAACCA 176
 Db 134760 ACTGTAATACACTGATGATTCGACACAGCTAATCTAAGATCAGCGCTGTGCA 134705

RESULT 14
 AC101994
 LOCUS Mus musculus clone RP24-406F14, WORKING DRAFT SEQUENCE, 4 unordered
 DEFINITION pieces.
 AC101994
 AC101994
 AC101994.2 GI:22381367
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 164972)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP24-406F14
 Unpublished
 TITLE
 REFERENCE 2 (bases 1 to 164972)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneu, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164972)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneu, L., Mihova, T., Mlenka, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060770.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L17843
 Center clone name: 406.F.14
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 164289 bases at least Q40
 Consensus quality: 164529 bases at least Q30
 Consensus quality: 164587 bases at least Q20
 Insert size: 151000; agarose-fp
 Insert size: 164672; sum-of-contigs
 Quality coverage: 9.7 in Q20 bases; agarose-fp
 Quality coverage: 8.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 150: contig of 50 bp in length
 * 51 150: gap of 100 bp
 * 151 27038: contig of 26888 bp in length
 * 27039 27138: gap of 100 bp
 * 27139 58986: contig of 31848 bp in length
 * 58987 59086: gap of 100 bp
 * 59087 164972: contig of 105886 bp in length.

FEATURES

source

1..164972
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-406P14"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1..50
 /note="assembly_fragment"
 151..27038
 /note="assembly_fragment"
 27139..58986
 /note="assembly_fragment"
 59087..164972
 /note="assembly_fragment"
 50238 a 31480 c 31671 g 51283 t 300 others

BASE COUNT 50238 a 31480 c 31671 g 51283 t 300 others

ORIGIN

Query Match 10.8%; Score 38.2; DB 2; Length 164972;
 Best Local Similarity 52.9%; Pred. No. 6.1;
 Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 Qy 2 GGAATTCGAGCAGATGAAGAAGCAGCGCTTGATTAATTCATAAGTTATTGGAAGA 61
 Db 43652 GAGAATTCAGGTTGATGAGAGAACTGATGTCATTATTTATCATCTACCAAAA 43711
 Qy 62 AAAAGGATTTACCGTTCATTTCCGTTATTCACAAACGCGTGATTACGGAATTCCTCCAAAG 121
 Db 43712 AATATGTTTCTCATTCATTCACCTTTCCATGGAGATCTATATCAGATTCATCAGG 43771
 Qy 122 CCGTAAAGATTTACGTTAATTCGAAACAGAAATAA 156
 Db 43772 CCCCCAAGCATTCGCCACATCTAAACTAAAGAA 43806

RESULT 15

AP003597 342850 bp DNA linear BCT 28-NOV-2001
 LOCUS Nostoc sp. PCC 7120 DNA, complete genome, section 17/19.
 DEFINITION AP003597 BA000019
 ACCESSION AP003597.1 GI:17133729
 VERSION
 KEYWORDS

SOURCE

Nostoc sp. PCC 7120
 Nostoc sp. PCC 7120
 ORGANISM Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

REFERENCE

1
 AUTHORS Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S.,
 Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T.,
 Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,
 Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M.,
 Yasuda, M. and Tabata, S.

Complete genomic sequence of the filamentous nitrogen-fixing

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

Cyanobacterium Anabaena sp. strain PCC 7120

1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail: kaneko@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/cyanobase/,
 Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)

FEATURES

source

1..342850
 /organism="Nostoc sp. PCC 7120"
 /mol_type="genomic DNA"
 /db_xref="taxon:103690"
 /note="synonym: Anabaena sp. PCC7120"
 complement (174..1478)
 /gene="all4593"
 complement (174..1478)
 /gene="all4593"
 /note="ORF_ID: all4593
 hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76292.1"
 /db_xref="GI:17133730"
 /translation="MDKSLIQLVDELPSDNTVVKVNLALDYVAPQWQNLVGFQDNTIR
 AVTGETDARVQIKINDRATSLYNDPQQYQSAVKLYQITDKADTAMATAALANKYGEK
 IGFLSFLNITPKADVTQSIDLVSKIAIEITVFKLNGIPQNPQEPFANSANNYQDA
 SLIRNALVCLDGLPLGDFLSKIQIGSAENSITQNPVFLAVNSLPGSNPTDK
 LCFINQGNVAVQWNNLVTKGITPQSISSHLGNFIQIADNDLDFVAAFQDQTNFYF
 PHTGQSVARLLKKAHLVKEEIKOETTOPIDQSSSAVKSDTSYAVSKKIEVWDEE
 FEDWYQATIEKVODDPPFINVIGGSHNENWVGKDDIETATYASSDANGAYAVKKVKC
 WSDDEETWYATIEKIQNQFIHTGYDSSYDEWNSDDIS"
 1680..1880
 /gene="asr4594"
 1680..1880
 /gene="asr4594"
 /note="ORF_ID: asr4594
 hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76293.1"
 /db_xref="GI:17133731"
 /translation="MTAIPQKSPQVLLYLGEDNYFVWEVPSLPGCISQKTRREA
 LTNIEVLQDKGFTLPEYQFINI"
 1877..2770
 /gene="alr4595"
 1877..2770
 /gene="alr4595"
 /note="ORF_ID: alr4595
 unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76294.1"
 /db_xref="GI:17133732"
 /translation="MSKLVDRLLLCFTFVLGFTTTHNVOAESAPFYWEFINVDIA
 VQNGDMLVTETQKYNFIRDVNHLSRVIPLDNVDRITVSVSENGQLPSTSTG
 NHQLLSNEHQLQSPESHTEVLYKRVIGGLVNHVYAOVYWEAFSDRAKAPQOVR
 VEPPELTKIKIFOSFGVADIRVDTKTEATVKQIEPTGTLKIOVIFDRAQTKI
 ETSQWQSTQFFYERLVLLGLLVFGIFPLLIIIVSRNGTGTSSSIVGSSIDYGGGG
 DYGGDYGGDYGGGGGGGGG"
 3082..4473
 /gene="alr4596"
 3082..4473
 /gene="alr4596"
 /note="ORF_ID: alr4596"
 /codon_start=1
 /transl_table=11
 /product="malic enzyme"
 /protein_id="BAB76295.1"
 /db_xref="GI:17133733"
 /translation="NADITPNSFSVTLRLQIPNEVGLASITQAIASSGGNLGQIDL
 IEQTRQVSLRDLTVAASAEHAETIQAQKAIPIKLYVDVETNLHGGKISITSR
 IETKVSDDLAMATPFGVRCVATAQDSQVNTLITKONTVAIVDGGAVLGLNLGP
 HAALPYMEGKAMLFKEFGLDAPFICLATQDTEETIIQAVNKLQVSLADIRIVNGAGAV
 CEETIKRLDELIDIPVHDDQGTAVTLAALFNSKLKVLQKSLADIRIVNGAGAV
 ALARLLRKAGAEEKIWMCKSGIISTRTDLNEEKQEFVAVQAQTLAGAVQADVFIV
 SAPGVLTPMWKSMKADIVFAMANPIPEIQELISKDVAVIATGCRSDYPNQINNVL"

gene

CDS

```

PPGVFRGALOCRAQTITTTMYLEAASAIASLVNPSDLSREHIIPSVFDERVVTAVAA
VQRAAREGARG"
5270. .6853
/gene="alr4597"
CDS
5270. .6853
/gene="alr4597"
/note="ORF_ID:alr4597"
/codon_start=1
/transl_table=11
/product="type I restriction-modification system DNA
methylase"
/protein_id="BAB76296.1"
/db_xref="GI:17133734"
/translation="MTBKAQKNSESEFQKWLKAAADKLKRNIDAAEYKHIVLGLIFL
KYSDAPEELHQKLLAGEGYAGANPEDREYSAENVFFVPEARWSYLGQAKQPD
KTVDLAMEAI EQNAKRLXGIIIPKVYQOQKLDQKSLGGIIGDITIGDABAQAD
VLGRVYVFLQOPALAEQKGGQFYTPESIVKLLVEMLPYNGRVDPCCGSGMFVQ
SEKFKVHQEGLDIDISYGOESNETTVYKLCRMNLARGIDGNSIKWNPBGSFLNDAHK
DLKADPVIAPPNDSPMGSELLRNDGRMLDKLVPVGNANFAMWYSHETYLAPTGS
AGFVLSNGSSNNTSGEGDIRKALVQKDLVDCIUMLPQTQFYNTGIPACLNFLSRYKN
GNKRDHRGEVLPFDASELGYMWRSSRAFTAEISKIADTYHEWKKSGGYSYKIDGF
CKSAIAEIEKHNFVLPFGRYVGIPDEVEDGVSFEKMSLTWALGCEQMRGQLDDE
IKQLLVGFIVDELVDVF"
6858. .7121
/gene="asr4598"
6858. .7121
/gene="asr4598"
/note="ORF_ID:asr4598"
/unknown_protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76297.1"
/db_xref="GI:17133735"
/translation="MNTQVILKLLNPVIEVKLHQAQSHQRTLEEEITSILEDVTEN
TPITTSKRDWSGFFFEQTCAGWGELLVREPOFEAQEREPLL"
7226. .7804
/gene="alr4599"
7226. .7804
/gene="alr4599"
/note="ORF_ID:alr4599"
/hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76298.1"
/db_xref="GI:17133736"
/translation="MTAVKCKFKPLTPPEYFIWEKKOQLPHEVLYNGEVYAMTGGTONH
GRASNIIFIVKGLHRCGGQCVGNSDCRINIFETKYVPDVSTCDERORTAIOAIQ
IFCLVIVLSPSTASYDRGRKFRUYRRNPSLDQDILLVDABKIAIDILYKRNKRNWEIF
NYQLGDNKILQSIDLSPFIESVYEDIVFEESY"
7855. .8145
/gene="asr4600"
7855. .8145
/gene="asr4600"
/note="ORF_ID:asr4600"
/unknown_protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76299.1"
/db_xref="GI:17133737"
/translation="MIFKKNMTQVILKQNLNPVLEKLELAQSHQRTLEEEITSILE
DVAKEEPEKSSSGFMDTLQFRKMQQENITFDADADIRDRSVGRDVEP"
8211. .8573
/gene="alr4601"
8211. .8573
/gene="alr4601"
/note="ORF_ID:alr4601"
/hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76300.1"
/db_xref="GI:17133738"
/translation="MVKKLTQHOLETATGVSVMHELLFGQLRLVESQKRRLLEYINQ
IPLKMTILNDLKAQWHAQERARLSKMGKTPAFIDGQIASIAYSNLLILVTNNVSDF

```

```

ESFNDLAVENMFVNSGEG"
8577. .9749
/gene="alr4602"
8577. .9749
/gene="alr4602"
/note="ORF_ID:alr4602"
/codon_start=1
/transl_table=11
/product="type I site-specific deoxyribonuclease chain S"
/protein_id="BAB76301.1"
/db_xref="GI:17133739"
/translation="MSEWKETTLGEIADIIMGQSPGTGCTCNNNGQGLPLNGPTFEGD
RNPFTQITIDFKKIAAGDGLFCVRGTTGRMNWADQKKAIGRGISIRAKDGLIFQ
PYIRATIEKESKSLAVATGTFPFIKSHLNLIVQLPSKNIKIYISNARIIDDKI
NLRSQNETLEAIATLTFKHWIFIDFEPNADGPKYSSGGAMVRSALGYIPEAWSVGK
LQVNLNKHGYAFKGEVITTEVTBKILLTPVNFKIGGQFNDKSKYYSADDYSNEYVL
RRKDLAITMOLSKEGDSLGYPAFIPDIKGVFLHNORIGKVENNNIDKTFLYELLOR
REYSHILGTSSGSTRVHTSPSRICEYSFVIPDFELIDKKSALATATIDKILLISTL"
9944. .10485
/gene="alr4603"
9944. .10485
/gene="alr4603"
/note="ORF_ID:alr4603"
/unknown_protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76302.1"
/db_xref="GI:17133740"
/translation="MIGKRLNDRNFSFHIAALEESQABDKYNLLISKVIDLLFV
KKLESISLIQYKNTLSNITDIRFLKFDGHSILKKNHDLIVDNNFIRTNIDNLNEYRLF"

Query Match          10.8%; Score 38.2; DB 1; Length 342850;
Best Local Similarity 55.7%; Pred. No.6.1;
Matches 73; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY      34 TTGATAATTTTCATAAAGTTATTGGAAGAAAAAGGATTACCGTCATTTCGGTATTCACA 93
Db      243329 TAGATATTATTAAAGAAAAACATTGGAAGAACTTGGTTATTTTGTGGAAGTATGGGTTTAA 243388

QY      94 ATACGGCTGATTACGGAAATTCCTCCAAAGCCCGTAAAGATTTACGTTAATTCGAACACAGAA 153
Db      243389 ATGCTGCGGAATATGGCAATCCGCAAAATTAGAGAACGTAATTTTATTGTTGGCAATAAA 243448

QY      154 TAACCAACAGAA 164
Db      243449 AAGGTAAAGTA 243459

Search completed: November 15, 2003, 03:34:00
Job time : 1567.62 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 227.827 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-83

Perfect score: 353

Sequence: 1 CGGAATCTGAGCAGATG.....ACGGAGGACCGTGGCTTT 353

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	353	19 AAV03591	Neisseria meningitidis
2	256.4	72.6	510	25 ABZ37757	N. gonorrhoeae nuc
3	256.4	72.6	510	25 ABZ41744	N. gonorrhoeae nuc
4	71	20.1	462	25 ABZ37756	N. gonorrhoeae nuc
5	71	20.1	462	25 ABZ41577	N. gonorrhoeae nuc
6	43.4	12.3	4519	19 AAV31192	E. coli J96 pathog
7	43.4	12.3	11165	21 ARA15186	DNA encoding Esche
8	35.6	10.1	13084	24 ABL34179	Human immune syste

C 9	35.4	10.0	935	21	AAZ80542	Human colon cancer
C 10	35.2	10.0	26518	22	AAH41227	Pyrococcus abyssi
C 11	34.2	9.7	1634	24	ABQ54282	Human ovarian anti
C 12	34	9.6	622	22	AAI17822	Human breast cancer
C 13	33.6	9.5	512	24	ABQ29518	Oligonucleotide fo
C 14	33.6	9.5	512	24	ABQ29519	Oligonucleotide fo
C 15	33.6	9.5	9742	24	ABL70479	Chemically treated
C 16	33	9.3	960	19	AAV53346	DNA encoding a Sta
C 17	33	9.3	2401	22	AAF6506	Arabidopsis CDC27A
C 18	33	9.3	2434	22	AAF6505	Arabidopsis CDC27A
C 19	32.6	9.2	14142	24	ABA01435	Streptococcus ther
C 20	32.4	9.2	421	22	AAK57045	Human immune/haema
C 21	32.4	9.2	628	21	AAI16500	Human colon cancer
C 22	32.2	9.1	2356	24	ABQ70511	Listeria monocytog
C 23	32.2	9.1	9155	24	ABD32463	Human immune syste
C 24	32	9.1	537	24	ABT10194	Human breast cancer
C 25	32	9.1	58073	18	AAT58840	Mycoplasma genital
C 26	32	9.1	64061	24	ABA92787	Buchnera sp. genom
C 27	31.8	9.0	1564	21	AAAC98092	Human colon cancer
C 28	31.8	9.0	2195	22	AAH15162	Human CDNA sequenc
C 29	31.8	9.0	7434	20	AAK12971	Enterococcus faeca
C 30	31.8	9.0	7434	24	ABS98766	Enterococcus faeca
C 31	31.6	9.0	283	24	ABN75274	Human nucleas-e-lik
C 32	31.6	9.0	951	24	ABK78842	Bacillus clausii g
C 33	31.6	9.0	2854	24	ABL34346	Human immune syste
C 34	31.6	9.0	3312	25	ABZ58867	Human PPIP-180 po
C 35	31.6	9.0	4204	25	ABX70993	Novel human CDNA s
C 36	31.6	9.0	6064	25	ABX17516	Human CDNA encodin
C 37	31.6	9.0	14388	23	ABL05670	Drosophila melanog
C 38	31.6	9.0	82952	24	ABN85766	Arabidopsis yellow
C 39	31.6	9.0	1664976	19	AAV21209	Methanococcus jann
C 40	31.4	8.9	14160	22	AAK65583	Human immune/haema
C 41	31.4	8.9	42738	22	AAK68992	Human immune/haema
C 42	31.2	8.8	883	23	ABV30371	Human prostate exp
C 43	31.2	8.8	1504	22	ABZ29913	C albicans apoptos
C 44	31.2	8.8	11384	24	ABK28222	DNA transcription
C 45	31.2	8.8	14924	24	ABL54322	Chemically treated

ALIGNMENTS

RESULT 1

AAV03591
ID AAV03591 standard; DNA: 353 BP.

AC AAV03591;

XX 22-OCT-1998 (first entry)

DT Neisseria meningitidis DNA sequence C130.

DE N. gonorrhoeae; N. lactamica; chromosome Z2491; region 1; region 2;

KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;

KW meningitis; ss.

XX Neisseria meningitidis.

OS Neisseria meningitidis.

PN WO9802547-A2.

XX 22-JAN-1998.

FD 11-JUL-1997; 97WO-FR01295.

XX 12-JUL-1996; 96PR-0008768.

PR (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (SMIK) SMITHKLINE BEECHAW.

XX Massif X, Tinsley C, Achtman M, Merker P, Ruelle J;

PI Vinals C;

XX

DR WPI; 1998-110594/10.
 XX Genes present in *Neisseria meningitidis* but not other *Neisseria*
 PT species - and related host cells, RNA, anti-sense sequences,
 PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*
 PT meningitidis infection and in protective vaccines
 XX
 PS Example 4; Pages 121-122; 150pp; French.
 XX
 CC AAV03575-606 represent sequences that are present in *Neisseria*
 CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the
 CC genes involved in biosynthesis of the capsule polysaccharide, *firpA* or *C*,
 CC *opc*, *porA*, *rotamase*, sequence IC1106, *IgA* protease, *pillin*, *pilC*,
 CC proteins which bind transferrin and opacity proteins. The DNA sequences
 CC are responsible for the differences in pathogenicity between *N.*
 CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that
 CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences
 CC *lactamica*, are responsible for colonisation and penetration of the
 CC mucosa. The DNA sequences can be used to produce probes and primers, and
 CC antibodies produced against the encoded proteins are used in standard
 CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*
 CC infection, particularly meningitis.
 XX
 SQ Sequence 353 BP; 119 A; 63 C; 82 G; 89 T; 0 other;
 Query Match 100.0%; Score 353; DB 19; Length 353;
 Best Local Similarity 100.0%; Pred. No. 6.7e-98;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGAATTCGAGCAGATGAAGAAAGCAGGCTTGATAATTCATAAAGTTATTGGAAG 60
 DB 1 CGGGAATTCGAGCAGATGAAGAAAGCAGGCTTGATAATTCATAAAGTTATTGGAAG 60
 QY 61 AAAAAGGATTTACGTCATTCGGTATTCACATACGGCTGATTACGGAATTCGCCAAA 120
 DB 61 AAAAAGGATTTACGTCATTCGGTATTCACATACGGCTGATTACGGAATTCGCCAAA 120
 QY 121 GCCGTAAAGATTTACGTTAATTCGAAACAGAAATACCAAAGAAAGCTGGAACCACTCA 180
 DB 121 GCCGTAAAGATTTACGTTAATTCGAAACAGAAATACCAAAGAAAGCTGGAACCACTCA 180
 QY 181 AGTATTCGGGCAACCGCTTACGTTACCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
 DB 181 AGTATTCGGGCAACCGCTTACGTTACCGATGTTTGGGAATGGAATGGCTTTCCCAA 240
 QY 241 CATTATTCGAGACACCAAGACGAAATTTATGTCATAGCTGTCGGGAATTATCTG 300
 DB 241 CATTATTCGAGACACCAAGACGAAATTTATGTCATAGCTGTCGGGAATTATCTG 300
 QY 301 ATATCACTTGAACGATTTGGCTTGATCTACCTAAACCGGAGAACCGTTGGCTTT 353
 DB 301 ATATCACTTGAACGATTTGGCTTGATCTACCTAAACCGGAGAACCGTTGGCTTT 353

RESULT 2
 ABZ37757
 ID ABZ37757 standard; DNA; 510 BP.
 XX
 AC ABZ37757;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE *N. gonorrhoeae* nucleotide sequence SEQ ID 103.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 PN W0200279243-A2.
 XX
 FD 10-OCT-2002.
 XX

PF 12-FEB-2002; 2002WO-IB02069.
 XX
 PR 12-FEB-2001; 2001GB-0003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Pontana MR, Pizza M, Massignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR P-PSDB; ABP76787.
 XX
 PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
 PT medicament for treating or preventing *N. gonorrhoeae* infection -
 XX
 PS Claim 6; Page 189; 815pp; English.
 XX
 CC The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
 CC molecules of the invention.
 XX
 SQ Sequence 510 BP; 173 A; 80 C; 117 G; 140 T; 0 other;
 Query Match 72.6%; Score 256.4; DB 25; Length 510;
 Best Local Similarity 95.5%; Pred. No. 2.5e-68;
 Matches 236; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
 QY 1 CGGGAATTCGAGCAGATGAAGAAAGCAGGCTTGATAATTCATAAAGTTATTGGAAG 60
 DB 203 CGGGAATTCGAGCAGATGAAGAAAGCAGGCTTGATAATTCATAAAGTTATTGGAAG 262
 QY 61 AAAAAGGATTTACGTCATTCGGTATTCACATACGGCTGATTACGGAATTCGCCAAA 120
 DB 263 AAAAAGGATTTACGTCATTCGGTATTCACATACGGCTGATTACGGAATTCGCCAAA 322
 QY 121 GCCGTAAAGATTTACGTTAATTCGAAACAGAAATACCAAAGAAAGCTGGAACCACTCA 180
 DB 323 GCCGTAAAGATTTACGTTAATTCGAAACAGAAATACCAAAGCTGGAACCACTCA 381
 QY 181 AGTATTCGGGCAACCGCTTACGTTACCGATGTTTGGGAATGGAATGGCTTTCCCAA 239
 DB 382 AGTATTCGGGCAACCGCTTACGTTACCGATGTTTGGGAATGGAATGGCTTTCCCAA 441
 QY 240 ACATTTATTCGAGACACCAAGACGAAATTTATGTCATAGCTGTCGGGAATTATCT 299
 DB 442 AAATTTATTCGAGACACCAAGACGAAATTTATGTCATAGCTGTCGGGAATTATCT 500
 QY 300 GATATCACTT 309
 DB 501 GATATCACTT 510

RESULT 3
 ABZ41744
 ID ABZ41744 standard; DNA; 510 BP.
 XX
 AC ABZ41744;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE *N. gonorrhoeae* nucleotide sequence SEQ ID 8077.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 PN W0200279243-A2.
 XX
 FD 10-OCT-2002.
 XX

XX 12-FEB-2002; 2002WO-IB02069.
XX 12-FEB-2001; 2001GB-0003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX P-PSDB; ABP80774.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection -
XX Disclosure; Page 778; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records AB237706-AB242016 represent nucleic acid
XX molecules of the invention.
XX Sequence 510 BP; 173 A; 80 C; 117 G; 140 T; 0 other;
SQ
Query Match 72.6%; Score 256.4; DB 25; Length 510;
Best Local Similarity 95.5%; Pred. No. 2.5e-68;
Matches 296; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
QY 1 CGGGAATTCCTGAGCAGATGAAGAAGCAGCGCTTGATAATTCATAAAGTTATTGGAAG 60
DB 203 CGGGAATTCCTGAGCAGATGAAGAAGCAGCGCTTGATAATTCATAAAGTTATTGGAAG 262
QY 61 AAAAGAGATTACCGTCCATTCGGTATTCACATACGGCTGATACGGATTCCCAAA 120
DB 263 AAAAGAGATTACCGTCCATTCGGTATTCACATACGGCTGATACGGATTCCCAAA 322
QY 121 GCCGTAAAGATTACGTTAATTCGAAACAGAAATACCAAGAAAGCTGGAACCACTCA 180
DB 323 GCCGTAAAGATTACGTTAATTCGAAACAGAAATACCAAGAAAGCTGGAACCACTCA 381
QY 181 AGTATTCGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGG-AAATGGCTTTCCCA 239
DB 382 AGTATTCGGCAACCGCTTACGGTAGCCGATGTTTGGGAATGGAAACCGCTTTCCCA 441
QY 240 ACATATTGCGAGGACACCAAGCAAGAAACGGATTTTATGCATAGCTGTCGGGATTATCT 299
DB 442 AAATTATGGCAGGACATCAAGACGAGCGGATTTTATGCATAGTTGTGCGGG-ATTATCC 500
QY 300 GATATCACTT 309
DB 501 GATATCAATT 510
RESULT 4
ABZ37756
ID ABZ37756 standard; DNA; 462 BP.
XX
AC
XX ABZ37756;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 101.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
XX
DR P-PSDB; ABP80774.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection -
XX Disclosure; Page 778; 815pp; English.
XX
PS The present invention relates to proteins from Neisseria gonorrhoeae.
XX
CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records AB237706-AB242016 represent nucleic acid
XX molecules of the invention.
XX Sequence 510 BP; 173 A; 80 C; 117 G; 140 T; 0 other;
SQ

PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
XX
DR P-PSDB; ABP76786.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection -
XX Claim 6; Page 188; 815pp; English.
XX
PS The present invention relates to proteins from Neisseria gonorrhoeae.
XX
CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records AB237706-AB242016 represent nucleic acid
XX molecules of the invention.
XX
SQ Sequence 462 BP; 135 A; 89 C; 114 G; 124 T; 0 other;
Query Match 20.1%; Score 71; DB 25; Length 462;
Best Local Similarity 85.7%; Pred. No. 1.1e-11;
Matches 114; Conservative 0; Mismatches 15; Indels 4; Gaps 3;
QY 215 TTGGGAATGGAAATGCGCTTCCACATATTGCGAGCACCAAGACGAAACGGATT 274
DB 1 TTGGGAATGGAAATGCGCTTCCACATATTGCGAGCACCAAGACGAAACGGATT 60
QY 275 ATCCATAGCTGTGGGAATTTCTGATATCACTT--GAACGATTGGC-TTGATACCTAA 331
DB 61 ATCCATAGCTGTGGGG-ATTATCCGATATCAATTGAAACGATTGGCTTTGATACCTAA 119
QY 332 AAACGGAGGAAC 344
DB 120 AAACGGAGGAAC 132
RESULT 5
ABZ41577
ID ABZ41577 standard; DNA; 462 BP.
XX
AC ABZ41577;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 7743.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
DR WPI; 2003-058415/05.

DR P-PSDB; ABP80607.

XX New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a

PT medicament for treating or preventing *N. gonorrhoeae* infection -

XX

XX Disclosure; Page 754; 815pp; English.

XX The present invention relates to proteins from *Neisseria gonorrhoeae*.

CC Also disclosed are the nucleic acid molecules encoding the proteins and

CC antibodies that specifically bind to the proteins. The composition

CC comprising the protein, nucleic acid or antibody is useful for the

CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*

CC infection, this may be in the form of a vaccine or gene therapy.

CC Sequences given in records AB237706-AB242016 represent nucleic acid

CC molecules of the invention.

XX

SQ Sequence 462 BP; 135 A; 89 C; 114 G; 124 T; 0 other;

Query Match 20.1%; Score 71; DB 25; Length 462;

Best Local Similarity 85.7%; Pred. No. 1.1e-11;

Matches 114; Conservative 0; Mismatches 15; Indels 4; Gaps 3;

QY 215 TTGGGAATGGAATGCGTTTCCCAACATTTTCGCGACACCAAGACGAAACGGATTTT 274

DB 1 TTGGGAATGGAATGCGTTTCCCAACATTTTCGCGACACCAAGACGAAACGGATTTT 60

QY .275 ATGCATAGCTGTGGGGAATTTATCTGATATCACTT--GAACGATTTGGC-TTGATACCTAA 331

DB 61 ATGCATAGTTTGGCGG-ATTATCGGATATCAATTTGAACGATTTGCTTTGATACCTAA 119

QY 332 AAACGGAGGAACC 344

DB 120 AAACGGAGGAAC 132

RESULT 6

AAV31192/c

ID AAV31192 standard; DNA; 4519 BP.

XX

AC AAV31192;

XX

DT 01-OCT-1998 (first entry)

XX

DE *E. coli* J96 pathogenicity island contig #6.

XX

XX PAI; pathogenicity island; uropathogenic *E. coli* detection; PAI IV; pheR;

KW PAI V; pheV; vaccine; protective immune response; ds.

XX

OS *Escherichia coli*.

XX

XX WO9822575-A2.

PN

PD 28-MAY-1998.

XX

XX 21-NOV-1997; 97WO-US21347.

PF

XX 14-OCT-1997; 97US-0061953.

PR

XX 22-NOV-1996; 96US-0031626.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX (UYWI-) UNIV WISCONSIN.

PA

XX Choi GH, Dillon PJ, Welch RA;

PI

XX WPI; 1998-312461/27.

DR

XX New isolated uropathogenic *E. coli* nucleotide sequences - used to

PT develop products for the detection of pathogenic *E. coli* and to

PT elicit an immune response to pathogenic *E. coli*

XX

PS Claim 21; Page 94-96; 250pp; English.

XX

CC This sequence represents a *E. coli* strain J96 contig containing

CC pathogenicity island (PAI) sequences, and represents a nucleic acid

CC molecule of the invention. PAIs are large fragments of DNA which comprise

CC pathogenicity determinants. The sequences of the invention are taken from

CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)

CC on the *E. coli* chromosome and is greater than 170 kb. PAI V is located at

CC approximately 94 min (at pheR) on the *E. coli* chromosome and is

CC approximately 160 kb in size. Antibodies specific to the proteins encoded

CC by the PAI open reading frames of the invention can be used in kits to

CC detect uropathogenic *E. coli*. The proteins are used in vaccines to elicit

CC a protective immune response in an animal to the uropathogenic *E. coli*

CC strain J96.

XX

SQ Sequence 4519 BP; 1141 A; 1101 C; 918 G; 1352 T; 7 other;

Query Match 12.3%; Score 43.4; DB 19; Length 4519;

Best Local Similarity 55.7%; Pred. No. 0.0064;

Matches 83; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 21 AAAGAAGCAGCCTTCATAATTCATAAAGTTATTGGAGAAAGGATTTACCGTCCAT 80

DB 979 AAAGAAGCAGCCTTCAGAGTTTATTAAAGCTACTTAAAGAGTTAGAGTAACTATA 920

QY 81 TTCCGTATTACAATACGCTGATTACGGAATTTCCCAAGCCGTAAAGATTTACGTTA 140

DB 919 TCTTTTAGCCAAATGCTGAGAACTATGCGATTCCCAAGAAAGAAAGACTCGTGCTC 860

QY 141 ATTGCAACAGAAATACCAAGAAAGCT 169

DB 859 TTAGTAGTCGAGTAGTAGTAAAGTTACCCT 831

RESULT 7

AAAI5186

ID AAI5186 standard; DNA; 11165 BP.

XX

AC AAI5186;

XX

DT 04-SEP-2000 (first entry)

XX

DE DNA encoding *Escherichia coli* virulence proteins.

XX

XX Virulence protein; tatA; tatB; tatC; tatE; mdoG; creC; recG; yggN;

KW eck1; iroD; iroC; iroE; mtd2; msl; vaccine; infection;

KW Gram negative bacterium; ss.

XX

OS *Escherichia coli*.

XX

XX Key Location/Qualifiers

FT CDS 2..1099

FT /tag= a

FT /product= "virulence protein"

FT /note= "encodes AAY93237; no termination codon given"

FT CDS 1102..1488

FT /tag= b

FT /product= "virulence protein"

FT /note= "encodes AAY93238"

FT CDS 1573..1896

FT /tag= c

FT /product= "virulence protein"

FT /note= "encodes AAY93239"

FT CDS 1939..2196

FT /tag= d

FT /product= "virulence protein"

FT /note= "encodes AAY93240; no termination codon given"

FT CDS 2198..2533

FT /tag= e

FT /product= "virulence protein"

FT /note= "encodes AAY93241"

FT CDS 2613..3041

FT /tag= f

FT /product= "virulence protein"

FT /note= "encodes AAY93242; no termination codon given"

FT CDS 3054..3410

```

FT FT /*tag= g
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93243"
FT FT 3460..3705
FT FT /*tag= h
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93244"
FT FT 3791..4837
FT FT /*tag= i
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93245"
FT FT 4878..7802
FT FT /*tag= j
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93246"
FT FT 7816..9483
FT FT /*tag= k
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93247"
FT FT 9836..10084
FT FT /*tag= l
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93248"
FT FT 10134..10430
FT FT /*tag= m
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93249"
FT FT 10459..10779
FT FT /*tag= n
FT FT /product= "virulence protein"
FT FT /note= "encodes AAY93250"
FT FT
XX XX
PN W2000028038-A2.
XX XX
XX 18-MAY-2000.
XX XX
XX 09-NOV-1999; 99NO-GB03721.
XX XX
XX 09-NOV-1998; 98GB-0024569.
XX 09-NOV-1998; 98GB-0024570.
XX 17-DEC-1998; 98GB-0027814.
XX 17-DEC-1998; 98GB-0027815.
XX 17-DEC-1998; 98GB-0027816.
XX 17-DEC-1998; 98GB-0027818.
XX 13-JAN-1999; 99GB-0000708.
XX 13-JAN-1999; 99GB-0000710.
XX 13-JAN-1999; 99GB-0000711.
XX 28-JAN-1999; 99GB-0001915.
XX XX
XX (MICR-) MICROSCIENCE LTD.
XX XX
XX Crooke HR, Clarke EE, Everest PH, Dougan G, Holden DW, Shea JE;
XX Feldman RG;
XX XX
XX W21; 2000-376550/32.
XX P-PSDB; AAY93237, AAY93238, AAY93239, AAY93240, AAY93241, AAY93242,
XX AAY93243.
XX XX
XX Peptide encoded by an operon including genes from Escherichia coli for
XX screening potential drugs, detecting virulence and treating conditions
XX associated with infection by a Gram negative bacterium -
XX PS Disclosure; Page 83-101; 122pp; English.
XX XX
XX The present sequence encodes Escherichia coli virulence proteins
XX The specification describes virulence proteins which are encoded
XX by an operon including tatA, tatB, tatC, tatE, mdoG, creC, recG, yggN,
XX eck1, iroD, iroC, iroE, mtd2 or mal-16 genes obtained from Escherichia
XX coli K1. The virulence proteins and polynucleotides, and their vaccines
XX are useful for screening potential drugs, for the detection of virulence
XX and for treating or preventing conditions associated with infection by
XX a Gram negative bacterium particularly Escherichia coli.

```

SQ Sequence 11165 BP; 3592 A; 2087 C; 2459 G; 3028 T; 0 other;
 Query Match 12.3%; Score 43.4; DB 21; Length 11165;
 Best Local Similarity 55.7%; Pred. No. 0.0088;
 Matches 83; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 21 AAGAAGCAGCGTTGATAATTTTCATAAGCTTATTCGAGAAAAAGATTACCGTCCAT 80
 DB 4178 AAGAAGGCGCCATTTCAGGAGTTTATTAAGCTACTTAAAGAGTTAGAGTATAACTATATA 4237
 QY 81 TTCGGTATTCACAATACGGCTGATTACGGAATTCGCCAAAGCCGTAAAAAGATTACGTTA 140
 DB 4238 TCATTTATAGCCAACTGCTGAGAACTATGGGATTCCTCCCAAGAGAAAAAGACTCGTGCTC 4297
 QY 141 ATTGCAACAGATAACCCAAAGAAAAGCT 169
 DB 4298 TTAGCTAGTCGAGTAGGTTAAAGTTACCT 4326
 RESULT 8
 ID ABL34179/C
 ID ABL34179 standard; DNA, 13084 BP.
 XX ABL34179;
 XX AC
 XX AC
 XX AC
 DT DT
 XX XX
 DE DE
 XX XX
 KW Human; immune system associated gene SEQ ID NO: 2152.
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosolic; neutropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
 KW antihemetic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; cancer; eye disease; arteriosclerosis; anaemia;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 KW KW
 XX KW
 OS Homo sapiens.
 XX OS
 PN WO200200928-A2.
 XX PN
 XX PD
 PD 03-JAN-2002.
 XX PD
 XX 02-JUL-2001; 2001WO-EF07537.
 XX PF
 PF 30-JUN-2000; 2000DE-1032529.
 XX PR
 PR 01-SEP-2000; 2000DE-1043826.
 XX PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX PA
 XX Olek A, Piepenbrock C, Berlin K;
 XX PI
 PI WPI; 2002-130909/17.
 XX DR
 DR Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX PT
 XX Claim 1; SEQ ID NO 2152; 32pp + Sequence Listing; German.
 PS
 PS The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX CC
 SQ Sequence 13084 BP; 3743 A; 243 C; 2982 G; 6116 T; 0 other;
 Query Match 10.1%; Score 35.6; DB 24; Length 13084;

Best Local Similarity 54.6%; Pred. No. 2.3;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 37 ATAAATTTTCATTAAGTTATTGGACGAAAAGGATTTCACGTCCTTCGTTTCGATTCACAATA 96
DB 9929 AAATATTTCATTAATAATATACAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9870

QY 97 CGGCTGATTACGGAATTCCTCCCAAGCCGCTAAAGATTTCAGTTAAATTCGAAAACAGATAA 156
DB 9869 TCAATATATAACACACCCCTTAATAACCAAAACATAAACTTTATCTATCTCTAAATAA 9810

QY 157 CCAAGAAAA 166
DB 9809 CCAAAAAAAA 9800

RESULT 9
AAZ80542/c
ID AAZ80542 standard; cDNA; 935 BP.

XX AC AAZ80542;
XX XX 07-APR-2000 (first entry)
XX XX Human colon cancer cell line SW480 cDNA clone SEQ ID NO:626.
XX DE Human; gene expression product; diagnosis; tumour; colon cancer;
XX KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
XX KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
XX KW hyperplasia; ds.
XX OS Homo sapiens.
XX PN WO9964576-A2.
XX XX 16-DEC-1999.
XX XX 09-JUN-1999; 99WO-IB01062.
XX PR 10-JUN-1999; 98US-0089801.
XX XX (FARB) BAYER CORP.
XX PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
XX PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
XX PI Schlegel R;
XX WPI; 2000-087220/07.

XX Novel nucleic acids, used to develop products for the diagnosis and
XX PT treatment of disorders involving unwanted cell proliferation,
XX PT particularly cancers, especially colon cancer.
XX XX Claim 15; Page 379; 469pp; English.

XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
XX CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
XX CC cDNA clones can be used to generate antisense oligonucleotides which
XX CC can be used for antisense therapy. Methods and products from the present
XX CC invention can be used for identifying and/or classifying cancerous cells
XX CC present in a human tumour, particularly in solid tumours, e.g.
XX CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
XX CC can be used for developing agents for the diagnosis and treatment of
XX CC disorders involving unwanted cell proliferation, such as neoplasia,
XX CC dysplasia or hyperplasia.

XX SQ Sequence 935 BP; 216 A; 216 C; 236 G; 211 T; 56 other;
Query Match 10.0%; Score 35.4; DB 21; Length 935;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 129 AGATTTCAGTTAATTCGAAACAGATAACCAAGAAAGCTGGAACCAAGTCAAGTATTCG 188

DB 342 AAATTTTNCACCCCAAGGAAANCCAAAGGCTTAATGGACCCCAATTGATTTCAG 283

QY 189 GCACAACGGCTTACGTTAGCCGATGTTTGGGAAT 223
DB 282 GCCCAAGGCTTGGGTACCCGAGAGCTTCGGGATT 248

RESULT 10
AAH41227/c
ID AAH41227 standard; DNA; 265118 BP.

XX AC AAH41227;
XX XX 29-OCT-2001 (first entry)
XX XX Pyrococcus abyssi genomic fragment #6.
XX DE Hyperthermophilic archaeon; hyperthermophilic protein; ds.
XX KW Pyrococcus abyssi.
XX OS Pyrococcus abyssi.
XX FH Key Location/Qualifiers
XX FT misc_feature 1..49980
XX FT /*tag= a
XX FT /note= "This sequence overlaps with the 3' end of
XX FT AAH41226"

XX PN FR2792651-A1.
XX XX 27-OCT-2000.
XX XX 21-APR-1999; 99FR-0005034.
XX XX 21-APR-1999; 99FR-0005034.
XX XX (CNRS) CNRS CENT NAT RECH SCI.
XX PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX DR New nucleotide sequences isolated from Pyrococcus abyssi encode
XX XX proteins useful in industry.
XX XX Claim 1; Page 593-665; 1657pp; French.

XX CC The present invention relates to the genomic sequence of Pyrococcus
XX CC abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a
XX CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
XX CC vents. The present sequence is a fragment of the genomic sequence of P.
XX CC abyssi. The 5' end of this sequence overlaps with the 3' end of AAH41226.
XX CC The proteins of the present invention have various potential industrial
XX CC uses, since the proteins are stable at very high temperatures, some up to
XX CC 110 degrees centigrade.
XX CC Note: This patent is in the same patent family as WO2000085062, which
XX CC contains additional sequences as shown in AAB99132-AAB99143,
XX CC AAB75903-AAH75920 and AAG66436.

XX SQ Sequence 265118 BP; 75226 A; 61302 C; 54005 G; 74585 T; 0 other;
Query Match 10.0%; Score 35.2; DB 22; Length 265118;
Best Local Similarity 52.8%; Pred. No. 8.5;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 112 TTCCCAAGCCGTAAGATTTCAGTTAATTCGAAACAGATAACCAAGAAAGAGCTGG 171
DB 91600 TACCCCAAGGATTGCAAGGGCCATGGTAAATCTAACGAGGCCAAGTGGAAATAG 91541
QY 172 AACCACTCAAGTATTCGGGCAACAGGCTTACCGTAGCCGATGTTTGGAAATGGAATGG 231

Db 91540 ATCTTTCATGGGTACAGGGGAATGCTTATGAGGCTGGTTTAATGGGCTGAAAGTTT 91491
 QY 232 CTTTCCCAACATTATGCGAGGACA 255
 Db 91480 ATGGAATAGATATAAGGGAGGACA 91457

RESULT 11
 ABQ54282/c
 ID ABQ54282 standard; cDNA; 1634 BP.
 XX ABQ54282;
 XX
 XX 22-AUG-2002 (first entry)
 XX
 XX Human ovarian antigen HNEUA49 cDNA, SEQ ID NO:162.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200200677-A1.
 XX
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI; 2002-147878/19.
 XX
 XX P-PSDB; ABP41205.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 1; SEQ ID No 162; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 XX
 SQ Sequence 1634 BP; 393 A; 403 C; 351 G; 473 T; 14 other;
 Query Match 9.7%; Score 34.2; DB 24; Length 1634;
 Best Local Similarity 50.0%; Pred. No. 2.9;
 Matches 81; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 24 GAAAGCAGGCTTGATAATTTTCATAAGTTATTGGAGAGAAAGGATTTACCGTCCATTTC 83
 Db 1631 GNACGGTGGGTTTWTTTTTTTTTTTTTTTTTTTTTTTTAAAAAGGTGGACTTGAATGTT 1572
 QY 84 GGTAATTCACAAATACGGGTGATTACGGAATTCCTCCAAAGCGTAAAGATTTACGTTAATT 143
 Db 1571 TATTACAGGATGCTGCAAGATAGGAATTCACATAGAAATTAGAAACCTAGTCAGAGG 1512
 QY 144 GCAACAGAAATAACCAAGAAAAAGCTGGAAACCAAGTCAAGTAT 185
 Db 1511 ACAAGCTTCATACAGTATGTACAGTTGGAACTGTTCAAGTAT 1470

RESULT 12
 AAL17822/c
 ID AAL17822 standard; cDNA; 622 BP.
 XX
 XX AAL17822;
 XX
 XX 07-DEC-2001 (first entry)
 XX
 XX Human breast cancer expressed polynucleotide 10279.
 XX
 XX Human; breast cancer; cell marker; cytostatic; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO200151628-A2.
 XX
 XX 19-JUL-2001.
 XX
 XX 10-JAN-2001; 2001WO-US00798.
 XX
 XX 14-JAN-2000; 2000US-0176077.
 XX
 XX 14-MAR-2000; 2000US-0189167.
 XX
 XX 24-MAR-2000; 2000US-0192099.
 XX
 XX 29-MAR-2000; 2000US-0193480.
 XX
 XX 15-MAY-2000; 2000US-0205230.
 XX
 XX 09-JUN-2000; 2000US-0211315.
 XX
 XX 25-JUL-2000; 2000US-0220534.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 XX WPI; 2001-451856/48.
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer -
 PS Claim 1; Page 1832-1833; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and

CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity.
 XX
 SQ Sequence 622 BP; 177 A; 91 C; 77 G; 277 T; 0 other;
 Query Match 9.6%; Score 34; DB 22; Length 622;
 Best Local Similarity 50.0%; Pred. No. 2.4; Indels 0; Gaps 0;
 Matches 85; Conservative 0; Mismatches 85;
 QY 80 TTTCGGTATTCCACATACGGCTGATTACGGAAATTCGCCAAGCGGTAAAGATTTACGTT 139
 DB 493 TTTCGGTATAATAAGGACAATACGGAAGGATAAAACAAAAAATGTTTAAATC 424
 QY 140 AATTGCAACAGATTAACCAAGAAAGCTGGACCAAGTCAAGTATTCGGGGCAACGGCT 199
 DB 423 ATTTTATGAGAGAAACAGGTAAAGTTCAATTCATCAAGATGCGAGTTAGCCATA 364
 QY 200 TACGTACCGCATGTTTTCGGGAATGGAAATGGCTTCCCAACATTATGC 249
 DB 363 TTGACTAACCAACTCTGTGACAGCTTCTTCACTGCCAAATTAATGC 314

RESULT 13
 ABO29518/c
 ID ABO29518 standard; DNA; 512 BP.
 XX
 AC ABO29518;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 16109.
 XX
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX

PT Determining the degree of cytosine methylation in genomic DNA, useful
 for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders

CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABO29518-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 512 BP; 100 A; 44 C; 145 G; 223 T; 0 other;
 Query Match 9.5%; Score 33.6; DB 24; Length 512;
 Best Local Similarity 63.8%; Pred. No. 3;
 Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 131 ATTTACGTTAATTGCAACAGATAACCAAGAAAGTGGACCAAGTCAAGTATTCGG 190
 DB 100 AATTACGTTAATCGTAATAATAACAAATCACTAAATCTAAACAATTCATCATCGTT 41
 QY 191 CAAACGGCTTACGGTAGCCG 210
 DB 40 CGAACAACTACGTTTATCG 21

RESULT 14
 ABO29519
 ID ABO29519 standard; DNA; 512 BP.
 XX
 AC ABO29519;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 16110.
 XX
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX

PT Determining the degree of cytosine methylation in genomic DNA, useful
 for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of

CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

XX SQ Sequence 512 BP; 223 A; 145 C; 44 G; 100 T; 0 other;
 Query Match 9.5%; Score 33.6; DB 24; Length 512;
 Best Local Similarity 63.8%; Pred. No. 3;
 Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 131 ATTACGTTAATGCAACAGAAATACCAAGAAAGCTGGACAGTCAAGTATTCGGG 190
 Db 413 AATTACGTTAATCGTAATTAACAATCAACTAAATCTAAACAAATTCATCCGTT 472
 QY 191 CAACGGCTTACGTTAGCG 210
 Db 473 CGAACAACTACGTTTAATCG 492

RESULT 15

ABL70479/c
 ID ABL70479 standard; DNA; 9742 BP.

AC ABL70479;

DT 01-JUL-2002 (first entry)

XX Chemically treated cell signalling DNA sequence#185.

KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytostatic; ds.

XX Unidentified.

XX WO200202807-A2.

XX 10-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07471.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-154758/20.

XX Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signalling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signalling

PS Claim 1; SEQ ID NO 369; 24pp+sequence listing; English.

XX The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligonucleotides for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated

CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX SQ Sequence 9742 BP; 3039 A; 88 C; 1939 G; 4676 T; 0 other;

Query Match 9.5%; Score 33.6; DB 24; Length 9742;
 Best Local Similarity 52.9%; Pred. No. 8.3;
 Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 51 TTATTGGAAGAAAAGGATTTACGTCCTATTCGGTATTCACAATACGGCTGATTACGGA 110
 Db 5177 TTTATATAAAAAAAAATAATCAATAATAATTCCTAAAAACGAAACCGCAAAATATTAT 5118
 QY 111 ATTCCTCAAAACCGGTAAAGATTTACGTTAATTGCAACACAGATAACCAAGAAAGAGCTG 170
 Db 5117 AATCTCCAAAACGCTAAATAATAATAATAATCTACAAAATAACAAAAAAAACCTT 5058
 QY 171 GAACCACTCAAGTATT 186
 Db 5057 AATTAAATCACTTAAT 5042

Search completed: November 15, 2003, 00:35:33

Job time : 230.827 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 2354.89 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-83

Perfect score: 353

Sequence: 1 CGGGAATTCGACGAGATG.....ACGGAGGACCGTTGGCTTT 353

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.6	12.4	456	9	AA981355 vx56b03.r
2	39.8	11.3	498	28	AQ733611 HS 5558 A
3	39.6	11.2	1100	29	AL104520 Drosophil
4	39.6	11.2	1201	13	EX361080 BX361080

5	39.4	11.2	853	29	CNS00CWP
6	38.6	10.9	771	14	CB998090
7	37.6	10.7	1101	29	CNS0032P
8	36.8	10.4	208	13	BU091084
9	36.8	10.4	1198	23	CC252073
10	36.8	10.4	1593	29	CC210958
11	36.6	10.4	967	14	CB562106
12	36.6	10.4	1018	13	EX400132
13	36.6	10.4	1101	29	CNS00ZCP
14	36.4	10.3	754	29	EX146252
15	36.2	10.3	499	28	EX146252
16	36.2	10.3	752	29	EX182757
17	36.2	10.3	1101	29	CNS0182P
18	36.2	10.3	1146	13	EX446722
19	36	10.2	551	29	EX244802
20	36	10.2	1201	13	EX375994
21	35.8	10.1	483	14	CB403766
22	35.8	10.1	1101	29	CNS00FEW
23	35.6	10.1	775	29	EX149557
24	35.4	10.0	331	12	BI315336
25	35.4	10.0	415	14	CD562547
26	35.4	10.0	464	10	EG019760
27	35.4	10.0	537	29	AG060614
28	35.4	10.0	550	12	BI447490
29	35.4	10.0	579	9	AW782350
30	35.4	10.0	772	14	CB201154
31	35.4	10.0	857	14	CB756442
32	35.4	10.0	866	29	BZ805512
33	35.4	10.0	949	29	CNS015RN
34	35.4	10.0	3209	11	AK085090
35	35.2	10.0	1201	9	AL514415
36	35	9.9	358	10	BE551716
37	34.8	9.9	557	28	BH336859
38	34.8	9.9	982	29	CNS006W7
39	34.8	9.9	1190	13	EX440012
40	34.8	9.9	1201	13	EX420578
41	34.6	9.8	601	29	EX207956
42	34.6	9.8	1200	13	EX414597
43	34.4	9.7	448	12	EG957424
44	34.4	9.7	544	13	BQ397837
45	34.4	9.7	567	28	AQ334267

ALIGNMENTS

RESULT 1
AA981355
LOCUS
DEFINITION
VX56b03.r1 Stratagene mouse macrophage (H937306) Mus musculus cDNA
Clone IMAGE:1279181.5, similar to TR:030357 O30357 MODIFICATION
METHYLASE ;, mRNA sequence.
ACCESSION
AA981355
VERSION
AA981355.1 GI:3159891
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 456)
Marrs,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished
COMMENT
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (infoimage.llnl.gov) for further information.
 MG1:670981
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 440.
 Location/Qualifiers
 1..456
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:1279181"
 /tissue_type="macrophage"
 /dev_stage="WEHI-3 cell line"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse macrophage (#937306)"
 /note="Organ: blood; Vector: pBluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. WEHI-3 cell line. Average insert size: 1.5 kb;
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG
 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

FEATURES

source

BASE COUNT 208 a 41 c 69 g 138 t
 ORIGIN
 Query Match 12.4%; Score 43.6; DB 9; Length 456;
 Best Local Similarity 58.5%; Pred. No. 1.7;
 Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 37 ATAAATTCATAAGTTATTGGAAGAAAAGGATTACCGTCGATTCGGTATTCACAATA 96
 Db 242 AAAAAATATAATGACTATAAATAATATAGGATATCAGATAATGTCACCACTTTAAATA 301
 QY 97 CGGCTGATACCGAATTCGCCAAGCCGTAAGAATTACGTTAATTCGAACAGATAA 156
 Db 302 GTGCAGATTGAGCTGTCTCAACAGAAAAGAGTAATTTTATAGCAATAAATA 361
 QY 157 CCAAGAGAAA 166
 Db 362 ACAAATTA 371

RESULT 2
 A0733611/c
 LOCUS
 DEFINITION HS_5558_A1_C09_T7A_RPCI-11 Human Male BAC library Homo sapiens
 Genomic clone Plate=1134 Col=17 Row=E, genomic survey sequence.
 A0733611
 VERSION A0733611.1 GI:5505163
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 498)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 99380589
 10449764
 COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Resear h Genetics (info@resgen.com). BAC end Web Server:
 http://www.husc.washington.edu
 Plate: 1134 row: E column: 17
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 498.
 Location/Qualifiers
 1..498

FEATURES

source

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=1134 Col=17 Row=E"
 /sex="male"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 145 a 112 c 59 g 174 t
 ORIGIN
 Query Match 11.3%; Score 39.8; DB 28; Length 498;
 Best Local Similarity 52.0%; Pred. No. 15;
 Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
 QY 14 CAGATGAAGAAGCAGCGCTTGATAATTTCATAAGTTATTGGAAGAAAAGGATTAC 73
 Db 323 CAGTTTGAGAGATACACGCGTTTATGCTGTGTCAGATCAGCATGAAGAGATTAG 264
 QY 74 CTTCCATTTTCGTTATTCACAATACGCTGATTCGGAATTCGCCAAGCCGTAAGAATT 133
 Db 263 CGTCATTGTATAAGTAAGATTAGGCTAAGAATGGAAGCTTGGGGAACAAATATTT 204
 QY 134 TACGTTAATGCAACAGATTAACCAAGAAAGCTGGACCAAGTCAAGTA 184
 Db 203 TACTTGAATTTGTGATGATCAATACCAACACCTCTCTAGAGAAATA 153

RESULT 3

CNS014RI/c

LOCUS
 DEFINITION Drosophila melanogaster Genome survey sequence SP6 end of BAC
 BACN12G24 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL104520
 VERSION AL104520.1 GI:5616534
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1100)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
 library (DrosBAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelOBAC11.
 Location/Qualifiers
 1..1100
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN12G24"

```

FEATURES
  source
    1. .853
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /db_xref="taxon:7227"
      /clone="BACR26M06"
      /clone_lib="RPCI-98"
      /note="end : T7"
  BASE COUNT      280 a   162 c   158 g   187 t   56 others
  ORIGIN
    11.2%; Score 39.4; DB 29; Length 853;
    Best Local Similarity 48.3%; Pred. No.21;
    Matches 73; Conservative 11; Mismatches 67; Indels 0; Gaps 0;
    37 ATAAATTTTCATRAAGTTATTGGAGAAAAAGGATTTACCGTCGATTTGGTGATTTCACAATA 96

```

Db 461 AAATTTTAAATAAAWAAAAAAGAGAGATTATTTAACTGGGAGCAATAAA 520
 Qy 97 CGGCTGATTACGGATTCCCAAGCCGTAAGATTACGTTAAATTCGCAACAGATAA 156
 Db 521 ATATTAGTAAATAAAWAAAAAAGAGAGATTATTTAACTGGGAGCAATAAA 580
 Qy 157 CCAAGAAAGCTGGAACCAAGTCAAGTATTC 187
 Db 581 AAAAATAATTAATAAAATAAAACATC 611

RESULT 6
 LOCUS CB998090 771 bp mRNA linear EST 01-MAY-2003
 DEFINITION AGENCOURT 13890706 NIH_MGC 148 Homo sapiens cDNA clone
 IMAGE:30348468 5', mRNA sequence.
 VERSION CB998090
 KEYWORDS CB998090.1 GI:30292610
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 771)
 AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Stefan Hanson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM391 row: d column: 13
 High quality sequence stop: 587.
 Location/Qualifiers
 1..771

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30348468"
 /tissue_type="pre-eclamptic placenta"
 /lab_host="DH10B TonA"
 /clone_lib="NIH_MGC 148"
 /note="organ: placenta; Vector: pBluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Library is oligo-dT primed and
 directionally cloned using primer
 5'-TTTTTTTTTTTTTTT-3', size-selected for average insert
 size 2.3 kb and normalized to ROT 5. This is a primary
 library enriched for full-length clones and constructed
 using the Cap-trapper method (Carninci, in preparation).
 Library constructed by M. Brownstein (NIH/NHGRI,
 National Institutes of Health). Note: this is a NIH_MGC
 Library."
 227 a 162 c 197 g 185 t

BASE COUNT 227 a 162 c 197 g 185 t
 ORIGIN
 Query Match 10.9%; Score 38.6; DB 14; Length 771;
 Best Local Similarity 50.8%; Pred. No. 32;
 Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 Qy 80 TTTCGGTATTCACATACGGGTGATTACGGAATTCCTCCCAAGCCGTAAGATTACGTT 139
 Db 537 TGTCCGGTGTAAACAAAATGAGTTCCACTCAGCCACCTACAGCCAGAGAGATAGGA 596
 Qy 140 AATTCCAAACAGAAATCCAAAGAAAGCTGAACCTCAAGTATTCGGGCAACCGCT 199
 Db 597 AATTGTTAGGAAATCAGCACTTACATTAAGAAATTTGGTACACCCCGACAGTAC 656

Qy 200 TAGGCTAGCCGATGTTTGGGAATCGAAATGCTTCCCAACATTATTGCGAGCACCAA 259
 Db 657 CATTTGGGGCCCAATTTCTGGTTCGAAATGGTGACACCAATTCCTGGGAGCCCAA 716
 Qy 260 G 260
 Db 717 G 717

RESULT 7
 LOCUS CNS0032P/c 1101 bp DNA linear GSS 03-JUN-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
 BACR07F18 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL063779.1 GI:4941535
 VERSION GSS.
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
 BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
 - Web: www.genoscope.cns.fr)

COMMENT Determination of this BAC-and sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osogawa and
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 source
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR07F18"
 /clone_lib="RPCI-98"
 /note="end: T7"
 BASE COUNT 330 a 188 c 53 g 431 t 119 others
 ORIGIN
 Query Match 10.7%; Score 37.6; DB 29; Length 1101;
 Best Local Similarity 39.9%; Pred. No. 60;
 Matches 87; Conservative 26; Mismatches 105; Indels 0; Gaps 0;

Qy 11 GAGCAGAAATCAAGAAAGCAGGCTGATTAATTCATAAGTTATTGGAGAAAGAGATT 70
 Db 924 GAARADWAAAAAAGAAAAAAATWKTAAATTAATGAATTTTAAATAATTKRAWT 865
 Qy 71 TACCGTCCATTTCCGTTATTCAATACGGCTGATTCAGGAATTCCTCCCAAGCCGTAAG 130
 Db 864 AANAATRTAATTAAGTTTAAAGATTAATTTTGGAAAGATTTTAAAGGRTTAAAT 805
 Qy 131 ATTACGTTAATTCGCAACAGAAATACCAAGAAAGCTGGAAACCAAGTCAAGTATTCGG 190
 Db 804 ATKTTAAATTTTGAATAAATTTAGAAAAATAAATTTTGGGAAATTTTAAATTTTGG 745
 Qy 191 CAACCGCCTACGGTAGCCGATGTTTGGGAATGGAAA 228

7	TCTGACGAGAAATGAAGAAACAGCGCTTGATAATTTTCATAAAGTTATTGGAAGAAAAAG	66
15	TAGGAACACAGACAGAGAAAAATGAATGAAAGATGAATCCCAAGGTTTATGCAGAAGATT	74
67	GATTTACCGCTCCATTTCGGTATTCAATACGGCTGATTACGGATTCCCAAGCCGTA	126
75	CATTTCATCTACCTTTTCTCGCTCATATTTAAATGTTAAATAATCAAAAATGATATA	134
127	AAAGATTTACGTTAATTCGAAACAGATAACCAAGAAAA	166
135	AAAAATTAATTTTATCTCTTCAAAAAAATAAAAAA	174
CC252073	1198 bp	DNA
CH261-18M7	Sp6.1	CH261 Gallus genomic clone
CC252073		genomic survey sequence.
CC252073.1	GI:30588823	
GSS		Gallus gallus (chicken)
SOURCE		Gallus gallus
ORGANISM		Gallus gallus
REFERENCE		1 (bases 1 to 1198)
AUTHORS		Kremitski, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E., and Wilson, R.
TITLE		Gallus gallus BAC End Reads
JOURNAL		Unpublished
COMMENT		Contact: Richard K. Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wustl.edu Insert Length: 182000 Std Error: 0.00 Seq primer: Sp6 ATTAGTGACACTAG Class: BAC ends High quality sequence start: 32 High quality sequence stop: 772.
FEATURES		Location/Qualifiers
source		1..1198
		/organism="Gallus gallus"
		/mol_type="genomic DNA"
		/strain="Red Jungle Fowl"
		/db_xref="taxon:9031"
		/clone="CH261-18M7"
		/sex="female"
		/cell_line="UCD001, inbred 256"
		/clone_lib="CH261"
		/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac"
BASE COUNT	365 a	202 c 271 g 355 t
ORIGIN		
Query Match	10.4%;	Score 36.8; DB 29; Length 1198;
Best Local Similarity	50.6%;	Pred. No. 96;
Matches	89;	Conservative 0; Mismatches 87; Indels 0; Gaps 0;
86	TATTCACAAATACCGCTGATTACGGAATTCGCCAAGCCGTAAGATTACGTTAATGC	145
357	TGTACACAGTAGTATGTTTTTCTGCATCTTAAATACATGATATTTATTTCATTCAATGG	298
146	AAACAGAAATACCAAGAAAGAGCTGGACAGTCAAGTATTCGGGCAACGGCTTACGGT	205
297	ATTAACATTGCACACTGAAAGAGAGAGTCTCTTAACCTGATCTTTCTAATTTACAGC	238
206	AGCCGATGTTTGGGAATGGAATGGCTTTCCCAACATTTATTCAGGACACCAAGA	261
237	ATCTTCAGTTTCAGGGAATGGAAGTCTCTACTAATATTTATGATAAAGACTCATA	182

```

RESULT 10
CC210958/1 1593 bp DNA linear GSS 12-MAY-2003
LOCUS CH261-170K18_RM1.1 CH261 Gallus gallus genomic clone CH261-170K18,
DEFINITION genomic survey sequence.
ACCESSION CC210958
VERSION CC210958.1 GI:30529626
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
Warren,W., Graves,T., Mardis,E. and Wilson,R.
Gallus gallus BAC End Reads
unpublished
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TACGACTCACTATAGGGAGA
Class: BAC ends
High quality sequence start: 360
High quality sequence stop: 1045.
Location/Qualifiers
1..1593
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-170K18"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pTAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

BASE COUNT 390 a 219 c 304 g 385 t 295 others
ORIGIN
Query Match 10.4%; Score 36.8; DB 29; Length 1593;
Best Local Similarity 50.6%; Pred. No. 1e+02;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 86 TATTCAATACGGTGTATAGGAAATCCCAAGCGTAAAGATTTCGTTAATGTC 145
Db 632 TGTACACAGTATGATTTTGTGTCATCTTAATATACATTGAATATTTTATTCATTCATGG 573
QY 146 AATACAGTAATACCAAGAAAGCTGGAACCTCAAGTATTTCGGCAACGGCTTACGGT 205
Db 572 ATAACTTGCACATGAAAGAGAAAGTCTCTCAACCTGATCTTTCTAATTACAGC 513
QY 206 AGCCGATGTTTGGGAATGGAATGGCTTCCCAATATTTCGAGGACACCAAGA 261
Db 512 ATCTTCAGTTCAGGAATGGAAGTGTCTACTAATATTATGTAATAGACTCATA 457

RESULT 11
CB562106 967 bp mRNA linear EST 02-APR-2003
LOCUS AGENCOURT_13321331 NICHD XGC Tad1 Xenopus laevis cDNA clone
DEFINITION IMAGE:6675382 3', mRNA sequence.
ACCESSION CB562106
VERSION CB562106.1 GI:29481636
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 967)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Drs. Donald Brown and Liquean Cai
cDNA library Preparation: CLONTECH
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM311 row: b column: 21
High quality sequence start: 15
High quality sequence stop: 517.
Location/Qualifiers
1..967
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6879382"
/dev_stage="metamorphosis stage 53"
/clone_lib="NICHD_XGC_Tad1"
/notes="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGCCGACATG-df(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.6 kb
(range 0.9-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 347 a 166 c 186 g 268 t
ORIGIN
Query Match 10.4%; Score 36.6; DB 14; Length 967;
Best Local Similarity 56.1%; Pred. No. 1e+02;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 51 TTATTCGAGAAAGAGATTACCGTCCATTTCGGTATTTCACATACGCTCATTCGGA 110
Db 499 TTATTTCAATAAAGGAACCTTAACCCCTTTTACTGTGAAAGATTGGCTGTATCTT 558
QY 111 ATTCCTCCAAAGCGTAAAGAGATTACGTTAATTGCAACAGCAATACCAAGAAAGCTG 170
Db 559 GAGAGACACAGGTTTAAAGTAACTAGCTATATATAAAATAAAACCAAAAAAATG 618
QY 171 GAA 173
Db 619 GAA 621

RESULT 12
BX400132 1018 bp mRNA linear EST 13-MAY-2003
LOCUS BX400132 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ004YM23 5-PRIME, mRNA sequence.
ACCESSION BX400132
VERSION BX400132.1 GI:30618078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1018)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

```

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5199.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ004AG12QF1cluster=5199.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ004AG12QF1.
Location/Qualifiers

FEATURES
source

```
1..1018
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ004YM23"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      291 a 116 c 175 g 238 t 198 others
ORIGIN
```

```
Query Match      10.4%; Score 36.6; DB 13; Length 1018;
Best Local Similarity 27.0%; Pred. No. 1.1e+02;
Matches 80; Conservative 75; Mismatches 14; Indels 0; Gaps 0;

QY 19 TGAAGAAGCAGCGTTCGATATTCATAAGTTATTGGAAGAAAGAGTTACCGTCC 78
Db 717 TTTTAAAAAARGAARGMRVTTTTTAAARVGGGAGGGAAGAAWARATAWTGGK 776
QY 79 ATTTCGTTATTCACATACGCGTGATTACGGAATTCGCCAAGCGTAAAGATTACGT 138
Db 777 TTTTGTGAGGAAGAAHAAWTGGGGGGGGAATTTTGGRAAMMACVHMGGGTTWAAA 836
QY 139 TAATTGCAACAGATATACCAAGAAAGCTGGACCACTCAAGTATTCGGGCAACGGC 198
Db 837 MMAGGGAAGAAAATTTTWWGGKTTAAKKTAAAMARVKGDKWTWAKYDKKKKK 896
QY 199 TTACGATACCGATGTTTGGGAATGGAATGGGTTTCCCAACATTATTCAGACACCA 258
Db 897 TTKCMMAKKDKKKKDDTAAABAAAATTTATTATAAAATTTTAAAKKWCMM 956
QY 259 AGACGAACGATTTTATGCTAGCTGTGCGGGAATATCTGATCATCTTGAACG 314
Db 957 MMCMKKKVCANVXCMNMCMNMCMKKKBDGDAAMMAAMMAKMKKKKTCB 1012
```

```
RESULT 13
LOCUS      CNS002CP/c
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN01312 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL097507
VERSION     AL097507.1 GI:5609118
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 1101)
REFERENCE   Genoscope.
            Direct Submission
            Submitted (23-JUL-1999) Genoscope - Centre National de Sequence :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
```

http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source

```
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN01E12"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"
BASE COUNT      346 a 170 c 129 g 306 t 150 others
ORIGIN
```

```
Query Match      10.4%; Score 36.6; DB 29; Length 1101;
Best Local Similarity 39.3%; Pred. No. 1.1e+02;
Matches 64; Conservative 28; Mismatches 71; Indels 0; Gaps 0;

QY 3 GGAATTCGAGCAGATGAAGAAAGCAGCGTTGATTAATTCATAAGTATTGGAAGAA 62
Db 1090 KDRWAAAWKTRWAAAAAFAAFAAFAAFAAFAAFAAFAAFAAFAAFAAFAAFAA 1031
QY 63 AAAGGATTTACCGTCCATTCGCTATTTCACATACGCGTACGGAATTCGCCAAGC 122
Db 1030 AAAATTTTWAAGTTTTTKKTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 971
QY 123 CGTAAAGATTTACGTTAATTCGAAACAGATAACCAAGAA 165
Db 970 TAAAAAATAAATTCGGTTTAAWAAAAAFAAFAAFAAFAAFAAFAAFAAFAAFAAFAA 928
```

RESULT 14
EX146252/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```
EX146252      754 bp DNA linear GSS 13-MAR-2003
Danio rerio genomic clone DKEY-10312, genomic survey sequence.
EX146252
EX146252.1 GI:27977610
GSS.
Danio rerio (zebrafish)
```

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 754)
Humphray, S.J., Huckle, E. and Durham, J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Unpublished

COMMENT

This sequence was generated from the T7 end of BAC 10312. 10312 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/projects/D_rerio/.

FEATURES
source

```
1..754
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-10312"
/tissue_type="testis"
/note="vector pIndigoBAC-536"
BASE COUNT      234 a 156 c 129 g 235 t
ORIGIN
```

```
Query Match      10.3%; Score 36.4; DB 29; Length 754;
Best Local Similarity 51.9%; Pred. No. 1.1e+02;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 105 TAGGAATTCGCCAAGCGTAAAGATTACGTTAATTCGAAACAGATAACCAAGAA 164
```

Db 747 TACTGCATATCCAAATATCTTACAAATAGTAGTATCCATGTACCAAAACAAACAGAG 688
 Qy 165 AAGCTGGAACCAAGTCAAGTATTCGGCAAAAGCGCTTACGGTAGCCGATGTTTGGGAATG 224
 Db 687 AAGCTGGAGAAAAGGAGCTATCTGGAATATCTGATCATGTATATATACTTTTTAGGAGCT 628
 Qy 225 GAATGCGTTTCCCAACATATTTCAGGACACCAAGAC 262
 Db 627 GAAAAGAGCGACACCATGTTAGGCGTAGATTCTGAGCC 590

RESULT 15
 BH030890 499 bp DNA linear GSS 17-JUL-2001
 LOCUS RPCI-24-23914:TJ RPCI-24 Mus musculus genomic clone RPCI-24-23914,
 DEFINITION genomic survey sequence.
 ACCESSION BH030890
 VERSION BH030890.1 GI:14800053
 KEYWORDS GSS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 499)
 Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
 Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
 Russell,D., de Jong,P. and Fraser,C.M.
 Mouse BAC End Sequences from Library RPCI-24
 UNPUBLISHED
 Other GSSs: RPCI-24-23914.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 239 row: I column: 4
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..499
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-23914"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /clone_lib="RPCI-24"
 /note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
 RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
 library was cloned in the pTARBAC1 cloning vector at the
 BamH1 sites using MboI partially digested male C57BL/6J
 DNA."
 BASE COUNT 148 a 74 c 114 g 163 t
 ORIGIN

Query Match 10.3%; Score 36.2; DB 28; Length 499;
 Best Local Similarity 50.3%; Pred. No. 1.2e+02;
 Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
 Qy 15 AGAATCAAGAGAGCGGCTTCATTAATTCATATAAGTTATTGGAGAAAAAGGATTACC 74
 Db 65 AGAGTGAGAGAGAGTGTGTGTGATGTGAGAAATAGTCAATCCCAAAAGAAACATGGT 124
 Qy 75 GTCCATTTCCGTTATTCACAAATACGGCTGATTACGGAAATCCCAAGCCGTAAGAATTT 134

Db 125 GAGTATTTTGCTATTTCACCTTACCAGGTGATTGGGTATTTACTGTAAATATAAACATAAGG 184
 Qy 135 ACGTTAATTGCAACAGAGATTAACCAAGAAAGCTGGACCAAGTCAAGTATTTCGGGC 191
 Db 185 ATTTAAATAGTGCCTAGCACATAGACAAATACCATGTTAGCTGTTATTTTAAAGTC 241

Search completed: November 15, 2003, 08:02:35
 Job time : 2358.89 secs


```
Db 979 AAGAAGGCCATTTTCAGGAGTTTATTAAAGCTACTTAAAGAGTTAGAGTATAACTATATA 920
QY 81 TTCGGTATTCAATACACGCTGATTACGGAATTCGCCAAGCCCTAAAGATTACGTTA 140
Db 919 TCTTTTATAGCAATGCTGAGAACTATGGGATTCGCCAAGAAAGAAAGAACTCGTGCTC 860
QY 141 ATTGCAACAGAAATAACCAAGAAAAGCT 169
Db 859 TTAGTAGTCGAGTAGGTAAGATTACCT 831

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 10.2%; Score 36; DB 1; Length 7218;
Best Local Similarity 1.7%; Pred. No. 0.17;
Matches 6; Conservative 197; Mismatches 147; Indels 0; Gaps 0;

QY 2 GGAATTCGACGACGAATGAAGAAGCAGGCTTCATTAATTCATAAGTTATTGGAAGA 61
Db 1406 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1347
QY 62 AAAAGATTACCGTCCATTTTCGGTATTCAACAATACGCTGATTACGGAATTCGCCAAG 121
Db 1346 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1287
QY 122 CCGTAAAGATTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAAACGATGCA 181
```

```
Db 1286 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1227
QY 182 GTATTCGGGCAACGCTTACGTTAGCCGATGTTTGGGAATGGAATGGCTTTCCCAAC 241
Db 1226 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1167
QY 242 ATTATTGAGGACACCAAGACGAAACGGATTTTATGCATAGCTGTGCGGAATTATCTCA 301
Db 1166 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107
QY 302 TATCACTTGAACGATTGGCTTGATACCTAAACACGAGAACCGTTGGCT 351
Db 1106 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1057

RESULT 3
US-09-328-111-626/c
; Sequence 626, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adrian
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/098,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 626
; LENGTH: 935
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(935)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-626

Query Match 10.0%; Score 35.4; DB 3; Length 935;
Best Local Similarity 60.0%; Pred. No. 0.12;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 129 AGATTTACGTTAATTGCAACAGATAACCAAGAAAAGCTGGAACCAAGTCAAGTATTG 188
Db 342 AATTTTTCACCCCTCCCAAGGGAANCAAGGCTTAATTGGACCAATTCATTGAC 283
QY 189 GCAAAACGCTTACGTTAGCCGATGTTTGGGAAT 223
Db 282 GCCAAAGGCTTTGGGTACCCGAAGCTTCGGATT 248

RESULT 4
US-09-453-702B-158
; Sequence 158, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Burland, Valerie
; APPLICANT: Perna, Nicole T.
```

Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6345723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296,95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE DESCRIPTION: SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 17136
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-09-453-702B-158
Query Match 9.5%; Score 33.6; DB 4; Length 17136;
Best Local Similarity 57.7%; Pred. No. 1.3;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 42 TTCATAAGTATTGGAAGAAAGGATTACCGTCCATTTCGGTATTTCACAAATACGGCT 101
Db 13604 TTCCTCTGTTATGCCAGCAAGGATATACATACAAAAGCATTTCTTCT 13663
QY 102 GATTACGGAATCCCAAGCGGTAAGATTACGTTAATGC 145
Db 13664 CATGATAAATTTCCACGCGTTATGATGATTGATTGCAATTC 13707
RESULT 5
US-08-936-165A-46/c
; Sequence 46, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-46
Query Match 9.3%; Score 33; DB 4; Length 960;
Best Local Similarity 51.7%; Pred. No. 0.65;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 21 AAAGAAGCAGGCTGCATAATTTTCATAAGTATTTCGAAGAAAGGATTACCGTCCAT 80
Db 882 AAAGAAGTAATGATGTTGTTTTCGACGAGATAAATGAATCATCGGAATTCCTA 823
QY 81 TTCGGTATTACAAATACGGCTGATTACGGAATTTCCCAAGCGTAAAGATTGTTAAAGTAAATTAATT 140
Db 822 TTCTGTATGCCITGAAATGCTAAATCGTAAATGATGATGTTAAAGTAAATTAATT 763
QY 141 ATTGCAACAGATAACCAAGAAA 165
Db 762 ATAGAAACATRAAATTCACAAAA 738
RESULT 6
US-09-385-982-505
; Sequence 505, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: I1
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 505
; LENGTH: 628
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(528)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-505

Query Match 9.2%; Score 32.4; DB 3; Length 628;
Best Local Similarity 52.8%; Pred. No. 0.85;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 147 AACGAAATACCAAGAAAGCTGGAACAGCTCAAGTATTCGGCAACAGCGTTACGGTA 206
DB 448 AAAAAGAGAGGAAATTAATATTCCTTTGGTTGAGNGATNATGAA 507
QY 207 GCCGATGTTTGGGAATGGAATGGCTTCCCAACATATTGCGAGCACCAAGACGAAA 266
DB 508 GGGGAAAATTTGGGCTTGAATTTTAAAGTAATACTTTTCTAATAAGGTAAGTCTAGA 567
QY 267 CGGAT 271
DB 568 AGAAT 572

RESULT 7
US-08-545-528D-1/c
Sequence 1, Application US/08545528D
Patent No. 6537773
GENERAL INFORMATION:
APPLICANT: Fraser et al.
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
Patent No. 6537773
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB193P1
CURRENT APPLICATION NUMBER: US/08/545.528D
PRIOR FILING DATE: 1995-10-19
PRIOR APPLICATION NUMBER: US 08/488,018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473,545
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 580073
TYPE: DNA
ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match 9.1%; Score 32; DB 4; Length 580073;
Best Local Similarity 54.2%; Pred. No. 16;
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 42 TTCATAAGTTATTGCAAGAAAGGATTTACCGTCCATTTCCGTTATTCACAAATACGGCT 101
DB 505245 TTAAAGAGTTATCTAAACAAAAGAAATTTACTGAGGATTTAGAAAGTAAACCTAAACC 505186
QY 102 GATTACGGAATTCGCCAAGCGTAAAGATTACGTTAATGCAACAGCAATACCAAA 161
DB 505185 GCTTTTAAATGCTTGTGAACGTTTATCATGATTAACAGATTATGATATTACCAA 505126

RESULT 8
US-09-107-532A-2880
Sequence 2880, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2880:
SEQUENCE CHARACTERISTICS:
LENGTH: 1353 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5) LOCATION 1...1353
SEQUENCE DESCRIPTION: SEQ ID NO: 2880:
US-09-107-532A-2880

Query Match 9.0%; Score 31.6; DB 4; Length 1353;
Best Local Similarity 54.5%; Pred. No. 2;
Matches 85; Conservative 0; Mismatches 69; Indels 2; Gaps 1;
QY 126 AAAAGATTACGTTAATTGCAACAGATAACCAAGAAAGCTGGAACCTCAAGTAT 185
DB 470 AAAATTATGGCGTTGTTAGAAAAATATACGAAAAAGAGAGAAATACGCAAGGAT 529
QY 186 TCGGCAACAGCTTACGGTAGCCGATGTTTGGGAATGGAATGCTTCCCAACATTA 245
DB 530 TTCCGCCAGATGCTGA--ATTTCCTTTAATTTATCGAAAAAGGCTTATTGCAAGCTTA 587
QY 246 TTGCAGGACCAACAGAACGGAATTTTATGCATA 281
DB 588 CTTCAGACACGAGGACGACGAGTTTCGTAAA 623

RESULT 9
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3


```
SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103958)..(103958)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664855)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 9.0%; Score 31.6; DB 4; Length 1664976;
Best Local Similarity 50.7%; Pred. No. 31;
Matches 76; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 17 AATGAAGAAGACGCGCTTGATATTTTCATTAAGTTATTGGAAGAAAAAGGATTTACCGT 76
DB 104475 AAGGAAGAAGAGATATTTTAAAAATTGAAGGATTTGAATTTAAAAAGGAGGATTTGATA 104416
QY 77 CCATTTCGGTTTCAATACGCGTATTACGGAATCCCCAAGCGGTAAAAAGATTAC 136
DB 104415 AATATCTTTATTTCTCCAGATCACTATCAATAGAAACGGAAGATTTGTTG 104356
QY 137 GTTATTCACACAGAAATACCAAGAAAA 166
DB 104355 AAAAACTTAACCAACAAATTCAGATGAGA 104326

RESULT 10
US-08-485-158A-6/c
; Sequence 6, Application US/08485158A
; Patent No. 5859328
; GENERAL INFORMATION:
; APPLICANT: Nasrallah, June B.
; APPLICANT: Nasrallah, Mikhail E.
; APPLICANT: Thorness, Mary K.
; TITLE OF INVENTION: ISOLATED DNA ELEMENTS THAT DIRECT
; TITLE OF INVENTION: PISTIL-SPECIFIC AND ANOTHER-SPECIFIC GENE EXPRESSION
; TITLE OF INVENTION: AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Sughrue, Mion, Zinn, Macpeak, & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.158A
; FILING DATE: 07-JUN-1995
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A-6217-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-485-158A-6
Query Match 8.8%; Score 31; DB 2; Length 475;
Best Local Similarity 52.8%; Pred. No. 2.1;
Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 39 AATTTCATTAAGTTATTGGAAGAAAAAGGATTTACCGTCCATTTCGGTATTACGAATACG 98
DB 160 ATTAGCTTAAATTCGTCGACTGAAGGAGAAAAAATTTAAATTCGTTTTTACTAAAAA 101
QY 99 GCTGATTACGGAATCCCCAAGCGGTAAAAAGATTTAGCTTAATTCGCAACAGATAACC 158
DB 100 GAGGGTTGAAAATTTGGTATTAACGTAATTTGGTTAACTATATATTAACACACCAAT 41
QY 159 AAGAAA 165
DB 40 AAAAAATA 34
RESULT 11
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 8.8%; Score 31; DB 4; Length 1830121;
Best Local Similarity 52.8%; Pred. No. 47;
Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 40 ATTTCATAAGTTATTGGAAGAAAAAGGATTTACCGTCCATTTCGGTATTTCGAATACG 99

Db 1105186 ATATTATTAATCAATTAATGTTTGGTTATGGTTGACCTCAGCTATATTAGATGCT 1105245
QY 100 CTGATTACGAATCCCAAGCCGTAAGAAATTTACGTTAATTCGAAACAGATAACCA 159
Db 1105246 CTACTGTGGTGTTCACAATCTAGAACAGATTTTCCTTAATTCGAAAATTAATAGTG 1105305
QY 160 AAGAAA 166
Db 1105306 AACATAA 1105312
RESULT 12
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
Query Match 8.8%; Score 31; DB 4; Length 1830121;
Best Local Similarity 52.8%; Pred. No. 47;
Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 40 ATTCATAGATTTTGGAGAAAAGGATTTACCGTCCATTCGGTATTCACAAATACGG 99
Db 1105186 ATATTATTAATCAATTAATGTTTGGTTATGGTTGACCTCAGCTATATTAGATGCT 1105245
QY 100 CTGATTACGAATCCCAAGCCGTAAGAAATTTACGTTAATTCGAAACAGATAACCA 159

Db 1105246 CTACTGTGGTGTTCACAATCTAGAACAGATTTTCCTTAATTCGAAAATTAATAGTG 1105305
QY 160 AAGAAA 166
Db 1105306 AACATAA 1105312
RESULT 13
US-09-734-674-3/c
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(202001)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3
Query Match 8.7%; Score 30.8; DB 4; Length 202001;
Best Local Similarity 54.4%; Pred. No. 25;
Matches 62; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 2 GGAATTCGAGCAGAAATGAAGAAAGCAGGCTTGATAATTTCAATAAGTTATTGAAGA 61
Db 3209 GGAAGTCTGACAATAAACAAATTACATTAAATGAATATATTTTCATGAGTAACTAGA 3150
QY 62 AAAGGATTTACCGTCCATTCGGTATTCACAATAGCGGTGATTCGGGATTC 115
Db 3149 AAAGCAAGCACCCCTACATCTTAATATATATCATATAATTCATGAGGTATACC 3096
RESULT 14
US-07-712-833A-1
; Sequence 1, Application US/07712833A
; Patent No. 5175101
; GENERAL INFORMATION:
; APPLICANT: GOTZ, Friedrich
; APPLICANT: SEEHER, Stefan
; TITLE OF INVENTION: RECOMBINANT RESTRICTION ENZYME SAU3AI
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Nikaido, Marmelstein, Kubovcik &
; ADDRESSEE: Murray
; STREET: 1725 K Street N.W., Suite 1000
; CITY: Washington D.C.
; COUNTRY: United States of America
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/712,833A
; FILING DATE: 19910610
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 4018441.2
; FILING DATE: 08-JUN-1990
; ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 199.897 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-83

Perfect score: 353

Sequence: 1 CGGGAATTCGAGCAGATG.....ACGGAGAACCGTTGGCTTT 353

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2169961 seqs, 134102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	353	10	US-09-928-457-83
2	43.4	12.3	4519	9	US-09-956-004-6
3	37.2	10.5	557	12	US-10-027-632-201983
4	37.2	10.5	557	12	US-10-027-632-201984
5	37.2	10.5	557	12	US-10-027-632-201985
6	37.2	10.5	557	12	US-10-027-632-201986
7	37.2	10.5	557	13	US-10-027-632-201987
8	37.2	10.5	557	13	US-10-027-632-201988
9	37.2	10.5	557	13	US-10-027-632-201989
10	37.2	10.5	557	13	US-10-027-632-201990
11	35.6	10.1	13084	12	US-10-027-632-201986
12	35.4	10.0	935	10	US-09-879-536-626
13	33.6	9.5	17136	14	US-10-114-170-158
14	33	9.3	960	9	US-09-939-980-46
15	33	9.3	2401	13	US-10-036-492-14
16	33	9.3	2434	13	US-10-036-492-9

Sequence 27, Appl
Sequence 505, App
Sequence 224164,
Sequence 224164,
Sequence 90983, A
Sequence 317596,
Sequence 90983, A
Sequence 317596,
Sequence 436, App
Sequence 7, Appl
Sequence 7, Appl
Sequence 87, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 102, App
Sequence 102, App
Sequence 34, Appl
Sequence 201070,
Sequence 220599,
Sequence 220600,
Sequence 220600,
Sequence 193599,
Sequence 193600,
Sequence 193601,
Sequence 193602,
Sequence 193599,
Sequence 193600,

Sequence 2512 13
Sequence 628 11
Sequence 528 13
Sequence 528 13
Sequence 558 12
Sequence 558 12
Sequence 558 12
Sequence 558 13
Sequence 9155 12
Sequence 10809 12
Sequence 10809 12
Sequence 271990 14
Sequence 580073 12
Sequence 640681 10
Sequence 1584 9
Sequence 1584 11
Sequence 7434 10
Sequence 598 12
Sequence 598 12
Sequence 598 12
Sequence 598 13
Sequence 657 12
Sequence 657 12
Sequence 657 12
Sequence 657 12
Sequence 657 13
Sequence 657 13

US-10-036-492-27
US-09-871-161-505
US-10-027-632-224164
US-10-027-632-224164
US-10-027-632-90983
US-10-027-632-317596
US-10-027-632-90983
US-10-027-632-317596
US-10-311-455-436
US-09-960-858-7
US-09-960-870-7
US-10-195-144-87
US-10-205-220-1
US-09-790-988-1
US-09-925-299-102
US-09-925-299-102
US-09-070-927A-34
US-10-027-632-201070
US-10-027-632-220599
US-10-027-632-220600
US-10-027-632-193599
US-10-027-632-193600
US-10-027-632-193601
US-10-027-632-193602
US-10-027-632-193599
US-10-027-632-193600

ALIGNMENTS

RESULT 1

US-928-457-83
Sequence 83, Application US/09928457
Patent No. US20020164603A1

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA, specific proteins and peptides
TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
TITLE OF INVENTION: for obtaining them and their biological application.

NUMBER OF SEQUENCES: 99

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30 (OEB)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/928,457

FILING DATE: 2001-08-14

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/214,759

FILING DATE: 199-12-10

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:

LENGTH: 353 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-928-457-83

Query Match 100.0%; Score 353; DB 10; Length 353;

Best Local Similarity 100.0%; Pred No. 5e-93; 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0;

1 CGGGAATTCGAGCAGATGAAGAAAGCAGCGCTTGAATTTCAATTAAGTTATTGGAAG 60

1 CGGGAATTCGAGCAGATGAAGAAAGCAGCGCTTGAATTTCAATTAAGTTATTGGAAG 60

QY 61 AAAAAGGATTTACCGTCATTTTCGGTATTCACAATACGGCTGATTACGGAATTTCCCAAA 120
 DB 61 AAAAAGGATTTACCGTCATTTTCGGTATTCACAATACGGCTGATTACGGAATTTCCCAAA 120
 QY 121 GCCGTAAAGATTTACGTTTAATTCGCAACAGATAAACAAGAAAGCTGGAAACCAAGTCA 180
 DB 121 GCCGTAAAGATTTACGTTTAATTCGCAACAGATAAACAAGAAAGCTGGAAACCAAGTCA 180
 QY 181 AGTATTCGGCAACCGCTTACGGTAGCCGATGTTTTGGGAATGGAAATGGCTTTCCCAA 240
 DB 181 AGTATTCGGCAACCGCTTACGGTAGCCGATGTTTTGGGAATGGAAATGGCTTTCCCAA 240
 QY 241 CATTATTCGAGACACCAAGACGAAACGGATTTTATCATAGCTGTGCGGGAATTTATCTG 300
 DB 241 CATTATTCGAGACACCAAGACGAAACGGATTTTATCATAGCTGTGCGGGAATTTATCTG 300
 QY 301 ATATCACTTGAACGATTGGCTTGATCTAAACACGGAGAACCGTTGGCTTT 353
 DB 301 ATATCACTTGAACGATTGGCTTGATCTAAACACGGAGAACCGTTGGCTTT 353

RESULT 2

US-09-956-004-6/c
 ; Sequence 6, Application US/09956004
 ; Patent No. US20020072595A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Patrick J. Dillon et al.
 ; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
 ; FILE REFERENCE: PB324D1
 ; CURRENT APPLICATION NUMBER: US/09/956,004
 ; CURRENT FILING DATE: 2001-09-20
 ; PRIOR APPLICATION NUMBER: 08/976,259
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/061,953
 ; PRIOR FILING DATE: 1997-10-14
 ; PRIOR APPLICATION NUMBER: 60/031,626
 ; PRIOR FILING DATE: 1996-11-22
 ; NUMBER OF SEQ ID NOS: 142
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 4519.

TYPE: DNA
 ORGANISM: Escherichia coli
 FEATURE:
 NAME/KEY: (3483)..(3483)
 LOCATION: (3483)..(3483)
 OTHER INFORMATION: n equals a, t, g, or c
 NAME/KEY: misc_feature
 LOCATION: (3487)..(3487)
 OTHER INFORMATION: n equals a, t, g, or c
 NAME/KEY: misc_feature
 LOCATION: (4292)..(4292)
 OTHER INFORMATION: n equals a, t, g, or c
 NAME/KEY: misc_feature
 LOCATION: (4318)..(4318)
 OTHER INFORMATION: n equals a, t, g, or c
 NAME/KEY: misc_feature
 LOCATION: (4329)..(4329)
 OTHER INFORMATION: n equals a, t, g, or c

US-09-956-004-6

Query Match 12.3%; Score 43.4; DB 9; Length 4519;
 Best Local Similarity 55.7%; Pred. No. 0.041; Mismatches 66; Indels 0; Gaps 0;
 Matches 83; Conservative 0
 QY 21 AAAGAAGCAGCGCTTCATAATTTCAATAAGTTATTGGAAGAAAAAGGATTTACCGTCCAT 80
 DB 979 AAAGAAGCAGCGCTTCATAATTTCAATAAGTTATTGGAAGAAAAAGGATTTACCGTCCAT 80
 QY 81 TTCGGTATTCACAATACCGCTGATTACGGAATTTCCCAAGCCGTAAGAAATTTACGTTA 140
 DB 919 TCTTTATAGCCAATGCTGAGAACTATGGGATTTCCCAAGAAAAAGAAAGAACTCGTGCTC 860

QY 141 ATTGCAACAGATAAACAAGAAAGCT 169
 DB 859 TTAGCTAGTCAGTAGGTAAAGTTACCT 831

RESULT 3

US-10-027-632-201983/c
 ; Sequence 201983, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1998-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 201983
 ; LENGTH: 557
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(557)
 ; OTHER INFORMATION: n = A,T,C or G

Query Match 10.5%; Score 37.2; DB 12; Length 557;
 Best Local Similarity 51.2%; Pred. No. 1;
 Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 25 AAAGCAGGCTTGATTAATTTCAATAAGTTATTGGAAGAAAAAGGATTTACCGTCCATTCG 84
 DB 406 AGAGAATGCCAATTAATTCGAAACATTAGTTTAAAGACTAGGTGGATACCTAGCATTCAT 347
 QY 85 GTATTCACAATACGGCTGATTACGGAATTCCTCCAAAGCCGTAAGAAATTTACGTTAATG 144
 DB 346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAAATTAATAAAACTTATTTTATTT 287
 QY 145 CAAACAGATAAACAAGAAAGCTGGAAACCAAGTCAAGTATTTCGGGCAAA 194
 DB 286 TAAACATAAAATAACATCATTCATGAATACAGTAGACTAATAGGACTAA 237

RESULT 4

US-10-027-632-201984/c
 ; Sequence 201984, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12

```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201984
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201984

Query Match
Best Local Similarity 10.5%; Score 37.2; DB 12; Length 557;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTCATATTTTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAGAATGCCAATTAATTGCAACATTAGTTTTAAGACTAGGTGGATACCTAGCATTCA 347
QY 85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAGCCGTAAGATTTACGTTAATTG 144
Db 346 GTAATGTTAATCTCTRACTTTAGAAAATGTCAAATTAATAAAGAACTATATTTATT 287
QY 145 CAAACAGATAACCAAGAAAAGCTGGAACAGTCAGTCAAGTATTCGGGCAA 194
Db 286 TAAACATAAAAATAACATCATTCATGAATACAGTAGACTAATAGGACTAA 237
```

```

RESULT 5
US-10-027-632-201985/c
; Sequence 201985, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 10827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201985
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201985

Query Match
Best Local Similarity 10.5%; Score 37.2; DB 12; Length 557;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTCATATTTTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAGAATGCCAATTAATTGCAACATTAGTTTTAAGACTAGGTGGATACCTAGCATTCA 347
QY 85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAGCCGTAAGATTTACGTTAATTG 144
Db 346 GTAATGTTAATCTCTRACTTTAGAAAATGTCAAATTAATAAAGAACTATATTTATT 287
QY 145 CAAACAGATAACCAAGAAAAGCTGGAACAGTCAGTCAAGTATTCGGGCAA 194
Db 286 TAAACATAAAAATAACATCATTCATGAATACAGTAGACTAATAGGACTAA 237
```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201985

Query Match
Best Local Similarity 10.5%; Score 37.2; DB 12; Length 557;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTCATATTTTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAGAATGCCAATTAATTGCAACATTAGTTTTAAGACTAGGTGGATACCTAGCATTCA 347
QY 85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAGCCGTAAGATTTACGTTAATTG 144
Db 346 GTAATGTTAATCTCTRACTTTAGAAAATGTCAAATTAATAAAGAACTATATTTATT 287
QY 145 CAAACAGATAACCAAGAAAAGCTGGAACAGTCAGTCAAGTATTCGGGCAA 194
Db 286 TAAACATAAAAATAACATCATTCATGAATACAGTAGACTAATAGGACTAA 237
```

```

RESULT 6
US-10-027-632-201986/c
; Sequence 201986, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 10827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201986
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201986

Query Match
Best Local Similarity 10.5%; Score 37.2; DB 12; Length 557;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTCATATTTTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAGAATGCCAATTAATTGCAACATTAGTTTTAAGACTAGGTGGATACCTAGCATTCA 347
QY 85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAGCCGTAAGATTTACGTTAATTG 144
Db 346 GTAATGTTAATCTCTRACTTTAGAAAATGTCAAATTAATAAAGAACTATATTTATT 287
QY 145 CAAACAGATAACCAAGAAAAGCTGGAACAGTCAGTCAAGTATTCGGGCAA 194
```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201986

Query Match
Best Local Similarity 10.5%; Score 37.2; DB 12; Length 557;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTCATATTTTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAGAATGCCAATTAATTGCAACATTAGTTTTAAGACTAGGTGGATACCTAGCATTCA 347
QY 85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAGCCGTAAGATTTACGTTAATTG 144
Db 346 GTAATGTTAATCTCTRACTTTAGAAAATGTCAAATTAATAAAGAACTATATTTATT 287
QY 145 CAAACAGATAACCAAGAAAAGCTGGAACAGTCAGTCAAGTATTCGGGCAA 194
```

```

Db      286 TAAACATAAAAATACATCATTCATGAATACAGTAGACTAATAGGACTAA 237
US-10-027-632-201983/c
; Sequence 201983, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201983
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201983
Query Match      10.5%; Score 37.2; DB 13; Length 557;
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      25 AAAGCAGCGCTTGATTAATTTCAATTAAGTTATTGGAGAAAAAGGATTTACCGTCCATTTCG 84
Db      406 AGAGAATGCCAATTAATTTGCAACACATTAGTTTAAAGACTAGGTGATACCTAGCATTCAT 347

QY      85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAAGCGGTAAAGAGATTTACGTTAATTG 144
Db      346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAATTTATAATAAAAACTATATTTTATT 287

QY      145 CAACAGAAATACCAAGAAAGCTGGAACCAAGTCAAGTATTCGGGCAA 194
Db      286 TAAACATAAAAATACATCATTCATGAATACAGTAGACTAATAGGACTAA 237

RESULT 8
US-10-027-632-201984/c
; Sequence 201984, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201984
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201985
Query Match      10.5%; Score 37.2; DB 13; Length 557;
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      25 AAAGCAGCGCTTGATTAATTTCAATTAAGTTATTGGAGAAAAAGGATTTACCGTCCATTTCG 84
Db      406 AGAGAATGCCAATTAATTTGCAACACATTAGTTTAAAGACTAGGTGATACCTAGCATTCAT 347

QY      85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAAGCGGTAAAGAGATTTACGTTAATTG 144
Db      346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAATTTATAATAAAAACTATATTTTATT 287

QY      145 CAACAGAAATACCAAGAAAGCTGGAACCAAGTCAAGTATTCGGGCAA 194
Db      286 TAAACATAAAAATACATCATTCATGAATACAGTAGACTAATAGGACTAA 237

RESULT 9
US-10-027-632-201985/c
; Sequence 201985, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201985
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201985
Query Match      10.5%; Score 37.2; DB 13; Length 557;
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      25 AAAGCAGCGCTTGATTAATTTCAATTAAGTTATTGGAGAAAAAGGATTTACCGTCCATTTCG 84
Db      406 AGAGAATGCCAATTAATTTGCAACACATTAGTTTAAAGACTAGGTGATACCTAGCATTCAT 347

QY      85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAAGCGGTAAAGAGATTTACGTTAATTG 144
Db      346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAATTTATAATAAAAACTATATTTTATT 287

QY      145 CAACAGAAATACCAAGAAAGCTGGAACCAAGTCAAGTATTCGGGCAA 194
Db      286 TAAACATAAAAATACATCATTCATGAATACAGTAGACTAATAGGACTAA 237

RESULT 8
US-10-027-632-201984/c
; Sequence 201984, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201984
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201985
Query Match      10.5%; Score 37.2; DB 13; Length 557;
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      25 AAAGCAGCGCTTGATTAATTTCAATTAAGTTATTGGAGAAAAAGGATTTACCGTCCATTTCG 84
Db      406 AGAGAATGCCAATTAATTTGCAACACATTAGTTTAAAGACTAGGTGATACCTAGCATTCAT 347

QY      85 GTATTCACAATACGGCTGATTACGGAATTCGCCAAAGCGGTAAAGAGATTTACGTTAATTG 144
Db      346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAATTTATAATAAAAACTATATTTTATT 287

QY      145 CAACAGAAATACCAAGAAAGCTGGAACCAAGTCAAGTATTCGGGCAA 194
Db      286 TAAACATAAAAATACATCATTCATGAATACAGTAGACTAATAGGACTAA 237

```



```
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTGATTAATTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAAGATGCCAATTAATTCGAACATAGTTTAAAGACTAGGTGATACCTAGCAATTCAT 347
QY 85 GTATTCAATACAGGCTGATTCAGGAATTCGCCAAGCGGTAAAGATTTACGTTAAATG 144
Db 346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAAATTAATAAATAAATATATTTTATT 287
QY 145 CAAACAGAAATAACCAAGAAAAGCTGGACCACTCAAGTATTTCGGGCAAA 194
Db 286 TAAACATAAAATAACATCATTCATGAATACAGTAGACTAATAGACTAA 237

RESULT 10
US-10-027-632-201986/c
; Sequence 201986, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201986
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(557)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-201986

Query Match 10.5%; Score 37.2; DB 13; Length 557;
Best Local Similarity 51.2%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 25 AAAGCAGGCTTGATTAATTCATAAAGTTATTGGAAGAAAAGGATTTACCGTCCATTTCG 84
Db 406 AGAAGATGCCAATTAATTCGAACATAGTTTAAAGACTAGGTGATACCTAGCAATTCAT 347
QY 85 GTATTCAATACAGGCTGATTCAGGAATTCGCCAAGCGGTAAAGATTTACGTTAAATG 144
Db 346 GTAATGTTAATCTCTACTTTAGAAAATGTCAAAATTAATAAATAAATATATTTTATT 287
QY 145 CAAACAGAAATAACCAAGAAAAGCTGGACCACTCAAGTATTTCGGGCAAA 194
Db 286 TAAACATAAAATAACATCATTCATGAATACAGTAGACTAATAGACTAA 237

RESULT 11
US-10-311-455-2152/c
; Sequence 2152, Application US/10311455
; Publication No. US20030143606A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
; FILE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2152
; LENGTH: 13084
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2152

Query Match 10.1%; Score 35.6; DB 12; Length 13084;
Best Local Similarity 54.6%; Pred. No. 12;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 37 ATAAATTTTCAATAAGTTATTGGAGAGAAAAGGATTTACCGTCCATTTCGGTATTTCACAATA 96
Db 9929 AATATTTTCAATAAATAATATACAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9870
QY 97 CGGCTGATTTACGGAATTCGCCAAGCGGTAAAGATTTACGTTAAATTCGAAACAGAAATAA 156
Db 9869 TCAATATATAACAAACCCGCTAAATAACCAAAACATAAACTTTATCTCTCTAAATAA 9810
QY 157 CCAAGAGAAA 166
Db 9809 CCAAAAAAAA 9800

RESULT 12
US-09-879-536-626/c
; Sequence 626, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 626
; LENGTH: 935
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

LOCATION: (1)...(935)
OTHER INFORMATION: n = A,T,C or G
US-09-879-536-626

Query Match 10.0%; Score 35.4; DB 10; Length 935;
Best Local Similarity 60.0%; Pred. No. 4.4;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 129 AGATTACGTTAAATGCAACAGAAATACCAAGAAAGCTGGAACAGTCAAGTATTCG 188
Db 342 AAATTTTNCACCCCAAGGAAANCCAAAGCTTAATTTGGACCAATTTTACG 283

QY 189 GCGAAACGGCTTACGGTAGCGATGTTTGGGAAT 223
Db 282 GCCCAAGGCTTTGGTACCCGAGCTTCGGGAT 248

RESULT 13
US-10-114-170-158
Sequence 158, Application US/10114170
Publication No. US20030023075A1
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27396
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 17136
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 158:

US-10-114-170-158
Query Match 9.5%; Score 33.6; DB 14; Length 17136;
Best Local Similarity 57.7%; Pred. No. 53;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 42 TTCATAAGCTTATGCAAGAAAGGATTTACCGTCCATTTGGTATTCACAAATACGGCT 101

Db 13604 TTCTTCCTGTTATTGCCAGCAAAAGGATATAACAATAACAAAAGCAATTGATTTCTTCT 13663

QY 102 GATTCAGGAATTCCTCCCAAGCGTAAAGATTTACGTTAATTCG 145

Db 13664 CATGATAAAATTTCCCAACGGCGTTAATGATGTTATTGAATTC 13707

RESULT 14
US-09-939-980-46/c
Sequence 46, Application US/09939980
Patent No. US20020082234A1
GENERAL INFORMATION:
APPLICANT: Black, Michael
Burnham, Martin
Hodgson, John
Knawles, David
Lonetto, Michael
Nicholas, Richard
Pratt, Julie
Reichard, Richard
Rosenberg, Martin
Ward, Judith
TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,980
FILING DATE: 27-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,165
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Glimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 46:

US-09-939-980-46
Query Match 9.3%; Score 33; DB 9; Length 960;
Best Local Similarity 51.7%; Pred. No. 22;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 21 AAGAAGCAGCGCTTGATTAATTTTATTAAGATTTTGCAGCAAGATTAATTCATCGGATTCCTA 80
Db 882 AAGAAGTAAATGATTTGTTTTCAGCAAGATTAATTAATTCATCGGATTCCTA 823

QY 81 TTCGGTATTCACAAATACGGCTGATTCAGGAATTCCTCCCAAGCGTAAAGATTTACGTTA 140

Db 822 TTGTGATGCTTGAATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAATGCTTA 763

Search completed: November 15, 2003, 08:32:04
Job time : 201.997 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 23:56:24 ; Search time 1195.2 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-88
Perfect score: 270
Sequence: 1 AATTTCGACGGGAGG.....CTCGAACTCAATTCATT 270

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 2045481396 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rtd.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_sy.*
39: em_hg_hum.*
40: em_hg_mus.*
41: em_hg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	270	100.0	270	6	A68917	A68917 Sequence 88
2	270	100.0	270	6	BD063017	BD063017 DNA and s
3	264.2	97.9	271	1	AF169464	AF169464 Neisseria
C 4	254.8	94.4	349061	1	NMA222491	AL162753 Neisseria
C 5	253.2	93.8	858	6	AX024111	AX024111 Sequence
C 6	251.6	93.2	858	6	AX024058	AX024058 Sequence
C 7	242.2	89.7	270	1	AF169428	AF169428 Neisseria
C 8	133.4	49.4	852	1	AF169459	AF169459 Neisseria
C 9	121.8	45.1	269	1	AF169451	AF169451 Neisseria
C 10	121.8	45.1	269	6	A68906	A68906 Sequence 77
C 11	121.8	45.1	269	6	BD063006	BD063006 DNA and s
C 12	48	17.8	267334	1	AE016997	AE016997 Chlamydo
C 13	44.6	16.5	3930	1	IM0302030	AJ302030 Listeria
C 14	43.4	16.1	92413	1	AL646075	AL646075 Ralstonia
C 15	43	15.9	4884	1	HGD1IM	X55141 H. Giganteus
C 16	42.6	15.8	2578	3	AY051762	AY051762 Drosophil
C 17	42.6	15.8	3170	3	AB049434	AB049434 Drosophil
C 18	42.6	15.8	20175	2	AC019990	AC019990 Drosophil
C 19	42.6	15.8	170498	3	AC008094	AC008094 Drosophil
C 20	42.6	15.8	244757	3	AE003678	AE003678 Drosophil
C 21	41.4	15.3	123098	9	AC008799	AC008799 Homo sapi
C 22	41.4	15.3	151031	9	AC008923	AC008923 Homo sapi
C 23	39.6	14.7	139746	2	AC079599	AC079599 Homo sapi
C 24	39.6	14.7	158531	9	AC025575	AC025575 Homo sapi
C 25	39.6	14.7	192773	2	AC092917	AC092917 Homo sapi
C 26	39.6	14.7	193010	2	AC041044	AC041044 Homo sapi
C 27	39.2	14.5	164845	2	AC121528	AC121528 Mus muscu
C 28	38.4	14.2	146087	2	AL845360	AL845360 Mus muscu
C 29	38.4	14.2	195985	2	EX119979	EX119979 Danio rer
C 30	38.4	14.2	213419	2	EX511083	EX511083 Danio rer
C 31	38.2	14.1	3276	5	BC044098	BC044098 Xenopus l
C 32	38.2	14.1	6068	6	AX344796	AX344796 Sequence
C 33	37.6	13.9	15455	1	AE002263	AE002263 Chlamydo
C 34	37.6	13.9	16903	1	AE001662	AE001662 Chlamydia
C 35	37.6	13.9	110000	6	AR310754	Continuation (10 o
C 36	37.6	13.9	181168	2	EX293540	EX293540 Danio rer
C 37	37.6	13.9	212999	2	AC111972	AC111972 Rattus no
C 38	37.6	13.9	274806	2	AC126582	AC126582 Rattus no
C 39	37.6	13.9	325865	1	AP025448	AP025448 Chlamydo
C 40	37.4	13.9	192199	2	AC139312	AC139312 Bos tauru
C 41	37.4	13.9	239901	2	AC125552	AC125552 Rattus no
C 42	37.4	13.9	262424	2	AC103010	AC103010 Rattus no
C 43	37.2	13.8	112063	9	AC091485	AC091485 Homo sapi
C 44	37.2	13.8	180488	2	AC125116	AC125116 Mus muscu

ALIGNMENTS

RESULT 1
A68917
LOCUS A68917 270 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 88 from Patent WO9802547.
ACCESSION A68917
VERSION A68917.1 GI:4759836
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 270)
AUTHORS Nassif, X., Tinsley, C., Achtman, M., Ruelle, J., Vinals, C. and
Merker, P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL

APPLICATIONS		/mol_type="genomic DNA"						
JOURNAL	Patent: WO 9802547-A 88 22-JAN-1998;	/db_xref="taxon:32644"						
COMMENT	INST NAT SANTE RECH MED (FR)							
FEATURES	Other publication FR 2751000 19980116.							
	Location/Qualifiers							
1..270								
/organism="unidentified"								
/mol_type="genomic DNA"								
/db_xref="taxon:32644"								
BASE COUNT	72 a 67 c 42 g 89 t							
ORIGIN								
Query Match 100.0%; Score 270; DB 6; Length 270;								
Best Local Similarity 100.0%; Pred. No. 1.1e-58;								
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
QY	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
DB	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
QY	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						
DB	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						
QY	121 CTCGGTTTATCGAATAACCTTAACCTTCACCTTCCTAGCACATGCATCGTAGGCATTCGC	180						
DB	121 CTCGGTTTATCGAATAACCTTAACCTTCACCTTCCTAGCACATGCATCGTAGGCATTCGC	180						
QY	181 TATCAACTCGGAATCGCAGGAACAGTGTGGAATACAAATCTTTACACCCAAATGTTTGA	240						
DB	181 TATCAACTCGGAATCGCAGGAACAGTGTGGAATACAAATCTTTACACCCAAATGTTTGA	240						
QY	241 TTACGGTTGGCTCGAACTCAATTTCAATT	270						
DB	241 TTACGGTTGGCTCGAACTCAATTTCAATT	270						
RESULT 3								
AF169464	271 bp DNA linear BCT 09-AUG-2000							
LOCUS	Neisseria meningitidis strain Z2491 clone Em024 unknown sequence.							
DEFINITION	AF169464							
ACCESSION	AF169464.1 GI:9754672							
VERSION								
KEYWORDS	Neisseria meningitidis							
SOURCE	Neisseria meningitidis							
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;							
	Neisseriaceae; Neisseria.							
REFERENCE	Perrin,A., Nassif,X. and Tinsley,C.R.							
AUTHORS	1 (bases 1 to 271)							
TITLE	Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae							
JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 271)							
AUTHORS	Perrin,A., Nassif,X. and Tinsley,C.R.							
TITLE	Direct Submission							
JOURNAL	Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156 rue de Vaugurard, Paris 75015, France							
FEATURES	Location/Qualifiers							
	1..271							
/organism="Neisseria meningitidis"								
/mol_type="genomic DNA"								
/strain="Z2491"								
/db_xref="taxon:487"								
/clone="Em024"								
BASE COUNT	72 a 67 c 42 g 89 t							
ORIGIN								
Query.Match 97.9%; Score 264.2; DB 1; Length 271;								
Best Local Similarity 98.5%; Pred. No. 3.4e-57;								
Matches 266; Conservative 0; Mismatches 4; Indels 0; Gaps 0;								
QY	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
DB	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
QY	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						
DB	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						

APPLICATIONS		/mol_type="genomic DNA"						
JOURNAL	Patent: WO 9802547-A 88 22-JAN-1998;	/db_xref="taxon:32644"						
COMMENT	INST NAT SANTE RECH MED (FR)							
FEATURES	Other publication FR 2751000 19980116.							
	Location/Qualifiers							
1..270								
/organism="unidentified"								
/mol_type="genomic DNA"								
/db_xref="taxon:32644"								
BASE COUNT	72 a 67 c 42 g 89 t							
ORIGIN								
Query Match 100.0%; Score 270; DB 6; Length 270;								
Best Local Similarity 100.0%; Pred. No. 1.1e-58;								
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
QY	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
DB	1 AATCTTCGGACGGGAGGCTTGTTTCTTCCTCTCTGTTCCGACCGGATTCTCAATA	60						
QY	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						
DB	61 AAAATCATTGATTCATCGAAGTTCAATCCATATACCATATCTTTAATAACGATTTTATG	120						
QY	121 CTCGGTTTATCGAATAACCTTAACCTTCCTAGCACATGCATCGTAGGCATTCGC	180						
DB	121 CTCGGTTTATCGAATAACCTTAACCTTCCTAGCACATGCATCGTAGGCATTCGC	180						
QY	181 TATCAACTCGGAATCGCAGGAACAGTGTGGAATACAAATCTTTACACCCAAATGTTTGA	240						
DB	181 TATCAACTCGGAATCGCAGGAACAGTGTGGAATACAAATCTTTACACCCAAATGTTTGA	240						
QY	241 TTACGGTTGGCTCGAACTCAATTTCAATT	270						
DB	241 TTACGGTTGGCTCGAACTCAATTTCAATT	270						
RESULT 2								
BD063017	270 bp DNA linear PAT 27-AUG-2002							
LOCUS	DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.							
DEFINITION	BD063017							
ACCESSION	BD063017.1 GI:22608620							
VERSION	JP 2001504684-A/79.							
KEYWORDS	unidentified							
SOURCE	unclassified.							
ORGANISM	Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.							
REFERENCE	1 (bases 1 to 270)							
AUTHORS	DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications							
TITLE	Patent: JP 2001504684-A 79 10-APR-2001;							
JOURNAL	INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM							
COMMENT	PN JP 2001504684-A/79							
PD 10-APR-2001								
PF 11-JUL-1997 JP 1998505685								
PR 12-JUL-1996 FR 96/08768								
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI CARLA VINALS,								
PI PETRA MERKER								
PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC								
Strandedness: Single;								
CC Topology: Linear;								
FEATURES	Location/Qualifiers.							
	1..270							
/organism="unidentified"								

Qy	121	CTCGGTTTATCGAATAACCTAACCTCCATCCATCCGTAGACATGATCGTAGGCATTGCG	180
Db	121	CTCGGTTTATCGAATAACCTAACCTCCATCCGTAGACATGATCGTAGGCATTGCG	180
Qy	181	TATCAACTCGGCAATCGCAGGAACAGTGTGCGAATACAAATCTTTACCCAAATGTTGGA	240
Db	181	TATCAACTCGGCAATCGCAGGAACAGTGTGCGAATACAAATCTTTACCCAAATGTTGGA	240
Qy	241	TTACGGTTGGCTCGAATCTCAATTTCAATT	270
Db	241	TTACGGTTGGCTCGAATCTCAATTTCAATT	270
RESULT 4			
LOCUS	NMA22491/c	349061 bp	DNA linear BCT 02-SEP-2002
DEFINITION	Neisseria meningitidis serogroup A strain 22491 complete genome;		
ACCESSION	segment 2/7.		
VERSION	AL162753 AL157959		
KEYWORDS	AL162753.2 GI:7379120		
SOURCE	Neisseria meningitidis Z2491		
ORGANISM	Neisseria meningitidis Z2491		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;		
AUTHORS	Neisseriaceae; Neisseria.		
TITLE	1 (bases 1 to 349061)		
	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,		
	Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,		
	Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,		
JOURNAL	Holroyd,S., Jagels,K., Leather,S., Moulé,S., Mungall,K.,		
	Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmons,M.,		
	Skellton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.		
	Complete DNA sequence of a serogroup A strain of Neisseria		
MEDLINE	meningitidis Z2491		
PUBMED	Nature 404 (6777), 502-506 (2000)		
REFERENCE	2 (bases 1 to 349061)		
AUTHORS	Parkhill,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria		
COMMENT	sequencing team, Sanger Centre, Wellcome Trust Genome Campus,		
	Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
	Notes:		
	Details of N. meningitidis sequencing at the Sanger Centre are		
FEATURES	available on the World Wide Web.		
	(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).		
	Location/Qualifiers		
	1. 349061		
repeat_unit	/organism="Neisseria meningitidis Z2491"		
	/mol_type="Genomic DNA"		
	/strain="Z2491"		
	/db_xref="taxon:122587"		
RBS	/note="serogroup: A"		
	complement(24..206)		
	/note="ATR repeat; hmms hit to HMM ATR (1 - 183), score:		
	310.39"		
gene	/label=ATR		
	209..212		
	220..681		
	/gene="NMA0368"		
CDS	220..681		
	/gene="NMA0368"		
	/note="NMA0368, probable integral membrane protein, len:		
	153 aa; contains four probable transmembrane domains"		
	/codon_start=1		
	/transl_table=11		
	/product="putative integral membrane protein"		
	/protein_id="CA583659.1"		
	/db_xref="GI:7379121"		
	/db_xref="SPTREMBL:Q9JWH7"		
	/translation="MQEQRKSSVFVFLIVSVALTIASLNVAFVLGHGSMGLT"		

```

/db_xref="GI:7379124"
/db_xref="SPTREMBL:Q9JWH4"
/translation="MDAVQSFERRHITTEIQSLQEDVIAGLNMAVELRQLDLOQA
QLRLYQKMDRNPDAQEPYSLRDEIPHY"
complement(3271..3423)
/gene="NMA0372"
complement(3271..3423)
/gene="NMA0372"
/notes="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=11
/product="very hypothetical protein NMA0372"
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="SPTREMBL:Q9JWH3"
/translation="MFCRLSPGSDGFLLAVYEPVGFICLQGTAYFRSGGRKPHRI
KGILSD"
misc_feature
3562..3571
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3586..4356)
/gene="thiF"
complement(3586..4356)
/gene="thiP"
/notes="NMA0373, thiP, probable ThiF protein, len: 256 aa;
similar to e.g. ThiF_ECOLI P30138 ThiF protein (251 aa),
fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopterin biosynthesis proteins). Contains
Pfam match to entry PF00899 ThiF_family, ThiF family"
/codon_start=1
/transl_table=11
/product="ThiF protein"
/protein_id="CAB83674.1"
/db_xref="GI:7379126"
/db_xref="SPTREMBL:Q9JWH2"
/translation="MTTEHNDNDAFLRYSRHILLDEIGEOOKLSAAHILVGGG
GLGAALPLAASGILTIADSDIVLHNLQROVAEDVGKIKTEALADRIRHN
HTDVATINELKDCRLTGLVQAADIVLDCNDYATQAVNACVQATPLVSGAAR
FEGVLAVRPLDPLSPYACILFDGGSADGICSLFVGFPLVIGLIGSTQAAEALKILL
DAGPFGHRLAVRALEGWQYFDLPNPECPVCGAER"
misc_feature
3787..3796
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3847..4257)
/gene="thiP"
/notes="Pfam match to entry PF00899 ThiF_family, ThiF
family, score 186.60, E-value 4.1e-52"
complement(4418..4427)
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4471..4724
/gene="ppc"
4471..4724
/gene="ppc"
/EC_numbers="4.1.1.31"
/notes="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
917 aa; similar to many e.g. CAP2_RHOA Q32483
phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
fasta scores; E(): 0, 43.3% identity in 928 aa overlap.
Contains 2x Pfam match to entry PF00311 PEPCase,
Phosphoenolpyruvate carboxylase, PS00017 ATP/Grp-binding
site motif A (p-loop); and PS00393 Phosphoenolpyruvate
carboxylase active site 2"
/codon_start=1
/transl_table=11
/product="phosphoenolpyruvate carboxylase"
/protein_id="CAB83675.1"
Query Match 94.4%; Score 254.8; DB 1; Length 349061;
Best Local Similarity 98.9%; Pred. No. 9.1e-55;

```

```

Matches 267; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 AATCTTCGCGACGGGAGGCTGTTTTCCTCCCTCTGTCCGACCGATTCTCAATA 60
Db 59503 AATCTTCGCGACGGGAGGCTGTTTTCCTCCCTCTGTCCGACCGATTCTCAATA 59444
QY 61 AAAATCATGTATTTCATCGAAGTTTCATTCCTATACCATTTCTTTAATAACGATTTTATG 120
Db 59443 AAAATCATGTATTTCATCGAAGTTTCATTCCTATACCATTTCTTTAATAACGATTTTATG 59384
QY 121 CTCGGTTTATCGAATAACCTAACTTCACCTTCGCTAGCACATGCTAGGCAATTCGC 180
Db 59383 CTCGGTTTATCGAATAACCTAACTTCACCTTCGCTAGCACATGCTAGGCAATTCGC 59324
QY 181 TATCAACTCGGCAATCGCAGGAACAGTGTGGGAATACAACTTTTACACCCAAATGTTTGA 240
Db 59323 TATCAACTCGGCAATCGCAGGAACAGTGTGGGAATAC-ATCTTTACACCCAAATGTTTGA 59265
QY 241 TTACGGTTGGCTCGAAACTCAATTTCAATT 270
Db 59264 TTACGGTTGGCTCGAAACTCAATTTCAATT 59235

RESULT 5
AX024111/c
LOCUS AX024111 858 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 54 from Patent FR2785293.
ACCESSION AX024111
VERSION AX024111.1 GI:10184422
KEYWORDS
SOURCE Neisseria gonorrhoeae
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
Bouchardon, A. and Renaud, M.G.
Patent: FR 2785293-A 54 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES
Location/Qualifiers
source
1..858
/organism="Neisseria gonorrhoeae"
/mol_type="genomic DNA"
/db_xref="taxon:485"
1..858
/notes="unnamed protein product"
CDS
/codon_start=1
/transl_table=11
/protein_id="CAC08931.1"
/db_xref="GI:10184423"
/translation="MSEELKMSFEPTVIEHLGVKMSHTVPAIAELIANAYDACATE
VEVRLFDKPEHKIVIKDNGIMSFDEINDFVLRGNRRREKQASPCGRITGRKGLG
KLALFRLGNKIEISTIQGNERTVFTLDVAEIKKSRIYQPEFQKESKRPENTENGTIT
LTFLPKQGYPLDNYVGHLSRLFDPAODPKIKVSLNGSEPRIIDNKLNLVTPQFE
WEYQDLATNISLSKSPQYEGYSGLIQGFITTEPKLNKMKGITLFANGRMNMPF
FIDSESSH"
BASE COUNT 302 a 151 c 178 g 227 t
ORIGIN
Query Match 93.8%; Score 253.2; DB 6; Length 858;
Best Local Similarity 98.5%; Pred. No. 2.2e-54;
Matches 266; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 AATCTTCGCGACGGGAGGCTGTTTTCCTCCCTCTGTCCGACCGATTCTCAATA 60
Db 282 AATCTTCGCGACGGGAGGCTGTTTTCCTCCCTCTGTCCGACCGATTCTCAATA 223
QY 61 AAAATCATGTATTTCATCGAAGTTTCATTCCTATACCATTTCTTTAATAACGATTTTATG 120
Db 222 AAAATCATGTATTTCATCGAAGTTTCATTCCTATACCATTTCTTTAATAACGATTTTATG 163
QY 121 CTCGGTTTATCGAATAACCTAACTTCACCTTCGCTAGCACATGCTAGGCAATTCGC 180

```


Db 162 CTCGGTTATCGAATAACCTTAACCTTCGATCGACACATGATCGTAGGCATTGCG 103

QY 181 TATCAACTCGGCATCGCAGGAACAGTGTGCGAATACAAATCTTTACACCCAAATGTTGGA 240

Db 102 TATCAACTCGGCATCGCAGGAACAGTGTGCGAATAC-ATCTTTACACCCAAATGTTGGA 44

QY 241 TTACGGTTGGCTCGAAACTCAATTTCAAAT 270

Db 43 TTACGGTTGGCTCGAAACTCAATTTCAAAT 14

RESULT 6					
AX024058/c					PAT 15-SEP-2000
LOCUS	AX024058	858 bp	DNA	linear	
DEFINITION	Sequence 1 from Patent FR2786293.				
ACCESSION	AX024058				
VERSION	AX024058.1	GI:10184370			

SOURCE	ORGANISM
Neisseria meningitidis	Neisseria meningitidis
Neisseria meningitidis	Bacteria; Proteobacteria; Neisseriales; Neisseriaceae; Neisseria.
Bacteria; Proteobacteria; Neisseriales; Neisseriaceae; Neisseria.	Neisseria meningitidis

REFERENCE
AUTHORS
JOURNAL

1 Nassif, X., Tinsley, C., Aujaime, L., Perrin, A., Rokbi, B., Bouchardon, A. and Renaud, M.G.
Patent: FR 2785293-A 1 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)

```

FEATURES
source
1.858
Location/Qualifiers
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"

```

```

CDS
1..858
/cw_xref=taxon:48/
/note="unnamed protein product"
/codon_start=1

```

```

/consort_start=1
/transl_table=1
/protein_id="CAC08905.1"
/db_xref="GI:10184371"
/translation="MSEELKMGFPETVIHLGVKMSVHTPAELIENAYDACATE
VEVRLDKPKHKIVIKDNGKMGFDEINDYFVILGRNREEKQACSGRPTGKGLIG
KLALFGLNKEIATISIOGNRVFTLDYAIRSKGIYQEPFKESVNESI ESGTIT
LTLETKQGVPLONYEENSRVTFDPAQDEKIKVSLNGSEPKIIDGNLKYDLVTPQFE
WEQDLDATNLISLSSKPEQYEGSLIQGKETTTEKPLKNMKMGITLFPANGRMVNMPEF
FTDSSESHF"

```

FIDESSHT			
BASE COUNT	297 a	145 c	182 g
ORIGIN			234 t
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

Query Match	93.2%	Score 251.6	DB 6	Length 858
Best Local Similarity	98.1%	Pred. NO. 5.7e-54		
Matches 265	Conservative	0	Mismatches 4	Indels 1

Qy 1 AATTTCITCCGCACGGGGAGGCTGTTTCTTCTCCCTTCTGTTCCGACCGATCTCAATA 60

Db 282 AATTTCITCCGCACGGGGAGGCTGTTTCTTCTCCCTTCTGTTCCGACCGATCTCAATA 223

QY	61	AAATCATTTGATTTCATCGAAGTTCATCCCTATACCGATTATCTTTAATAACGATTTTATG	120
Dδ	222	AAATCATTTGATTTCATCGAAGTTCATCCCTATACCGATTATCTTTGATAAAGATTTTATG	163

QY	121	CTCCGGTTTATCGAATAACCTAACTTCCACTCCGTAGCACATGCATCGTAGCATTCCG	180
Db	162	CTCCGGTTTATCGAATAACCTAACTTCCACTCCGTAGCACATGCATCGTAGCATTCCG	103

Qy 181 TATCAACTCGGCAATCGCAGGAAACAGTGTGGCAATCAATCTTTACACCCCAATGTTTCTGA 240

DB	102	TATCAACTCGGCAATCGCAGGACAGTGTGCGAATAC-ATCTTTACACCAAAATGTTCA	44
QY	241	TTACGGTTGGTCGGAACCACTCAATTTTCAATT	270

Db 43 TTACGGTTGGCTCGAACTCATTTTCAATT 14

1000

RESULT 7

AF169428
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE	AUTHORS	TITLE
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
SOURCE

BASE COUNT
ORIGIN

Query Match 89.7%; Score 242.2; DB 1; Length 270;
Best Local Similarity 98.2%; Pred. No. 1.4e-51;
Matches 266; Conservative 0; Mismatches 3; Indels 2

QY	1	AATCTTCGACGGGAGGCTTGTTTTCTT-CCCTTCTGTTCCGACCGATTCTCAAT	59
Db	1	AATCTTCGACGGGAGGCTTGTTTTCTTCTTCCCTTCTGTTCCGACCGATTCTCAAT	60
QY	60	AAAAATCATTCATTTCATCGAAGTTCATTCCTATACATTATCTTTAATAACGATTTTAT	119
Db	61	AAAAATCATTCATTTCATCGAAGTCATTTCCTATGTCATTCTTTTAAATAACGATTTTAT	120
QY	120	GCTCGGTTTTATCGAAATAAAGCTAACTTCACATTCCGTAGCACATGCATCGTAGGCATTTCG	179
Db	121	GCTCGGTTTTATCGAAATAAAGCTAACTTCACATTCCGTAGCACATGCATCGTAGGCATTTCG	180
QY	180	CTATCACTCGGGAATCGCAGGACAGTGTGCGAATACAACTTTTACACCCGAATGTTTCG	239
Db	181	CTATCACTCGGGAATCGCAGGACAGTGTGCGAATAC-ATCTTTACACCCGAATGTTTCG	239
QY	240	ATTACGGTTGGCTCGAAACTCAATTTCAATT	270
Db	240	ATTACGGTTGGCTCGAAACTCAATTTCAATT	270

RESULT 8
AF169459

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Accession number	Gene	Size (bp)	Strain	Reference
AF169459	Neisseria meningitidis	852	linear	BCT 09-AUG-2000
			clone 22491	Cml30 unknown sequence

AF169459
AF169459.1 GI:9754667
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacter
Neisseriaceae; Neisser

1. (bases 1 to 952)
Perrin, A., Nassif, X. and Tinsley, C.R.
Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to pathogenic *Neisseriae*
Unpublished

```

REFERENCE 2 (bases 1 to 852)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        1..852
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /strain="22491"
            /db_xref="taxon:487"
            /clone="Cm130"
BASE COUNT 272 a 154 c 206 g 220 t
ORIGIN
Query Match 49.4%; Score 133.4; DB 1; Length 852;
Best Local Similarity 98.6%; Pred. No. 9.2e-24;
Matches 145; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 124 CGGTTTATCGAATAACCTTAACCTTCCACTTCGGTAGCACATGCATCGTAGGCAATTCGCTAT 183
DB 852 CGGTTTATCGAATAACCTTAACCTTCCACTTCGGTAGCACATGCATCGTAGGCAATTCGCTAT 793
QY 184 CAATCTGGCAATCGCAGAAACAGTGTGGCAATACAACTTTACACCCAAATGTCGATTA 243
DB 792 CAATCTGGCAATCGCAGAAACAGTGTGGCAATAC-ATCTTTACACCCAAATGTCGATTA 734
QY 244 CGGTTGGTCCGAAACTCAATTTCAATT 270
DB 733 CGGTTGGTCCGAAACTCAATTTCAATT 707

RESULT 9
AF169451/c
LOCUS Neisseria meningitidis strain Z2491 clone Cm040 unknown sequence.
DEFINITION
ACCESSION AF169451
VERSION AF169451.1 GI:9754659
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 269)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae
JOURNAL Unpublished
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        1..269
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /strain="Z2491"
            /db_xref="taxon:487"
            /clone="Cm040"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
Query Match 45.1%; Score 121.8; DB 1; Length 269;
Best Local Similarity 98.4%; Pred. No. 8.4e-21;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATTCTTCGCACGGGAGGCTTGTTTCTCCCTTCCTGTCGACCGATTCTCAATA 60
DB 125 AATTCTTCGCACGGGAGGCTTGTTTCTCCCTTCCTGTCGACCGATTCTCAATA 66
QY 61 AAAATCATTGATTTCATCGAAGTTCATTCCTATACCATTCATTATTAACGATTTATG 120
DB 65 AAAATCATTGATTTCATCGAAGTTCGTCCTATACCATTCATTATTAACGATTTATG 6
QY 121 CTCGC 125
DB 5 CTCGC 1

RESULT 11
BD063006/c
LOCUS DNA
DEFINITION
ACCESSION BD063006
VERSION BD063006.1 GI:22608609
KEYWORDS JF 2001504684-A/89.
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Werker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 68 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,

```

```

DB 65 AAAATCATTGATTTCATCGAAGTTCGTCCTATACCATTCATTATTAACGATTTATG 6
QY 121 CTCGC 125
DB 5 CTCGC 1

A68906
Sequence 77 from Patent WO9802547.
A68906
ACCESSION A68906
VERSION A68906.1 GI:4759825
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Werker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
JOURNAL Patent: WO 9802547-A 77 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
COMMENT Other publication FR 2751000 19980116.
FEATURES
    source
        1..269
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
BASE COUNT 95 a 41 c 65 g 68 t
ORIGIN
Query Match 45.1%; Score 121.8; DB 6; Length 269;
Best Local Similarity 98.4%; Pred. No. 8.4e-21;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATTCTTCGCACGGGAGGCTTGTTTCTCCCTTCCTGTCGACCGATTCTCAATA 60
DB 125 AATTCTTCGCACGGGAGGCTTGTTTCTCCCTTCCTGTCGACCGATTCTCAATA 66
QY 61 AAAATCATTGATTTCATCGAAGTTCATTCCTATACCATTCATTATTAACGATTTATG 120
DB 65 AAAATCATTGATTTCATCGAAGTTCGTCCTATACCATTCATTATTAACGATTTATG 6
QY 121 CTCGC 125
DB 5 CTCGC 1

BD063006
Sequence 269 bp DNA linear PAT 27-AUG-2002
DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications.
ACCESSION BD063006
VERSION BD063006.1 GI:22608609
KEYWORDS JF 2001504684-A/89.
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 269)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Werker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 68 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,

```

COMMENT	SMITHKLINE BEECHAM		gene	
	PN	JP 2001504684-A/68	CDS	
	PD	10-APR-2001		
	PF	11-JUL-1997 JP 1998505685		
	PR	12-JUL-1996 FR 96/08768		
	PI	XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELE, PI CARLA VINALS,		
	PI	PETRA MERKER		
	PC	C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC		
	CC	Strandedness: Single;		
	CC	Topology: Linear;		
FEATURES	FH Key Location/Qualifiers.		gene	
	1..269		CDS	
	/organism="unidentified"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:32644"			
	95 a 41 c 65 g 68 t			
	BASE COUNT			
	ORIGIN			
	Query Match 45.1%; Score 121.8; DB 6; Length 269;			
	Best Local Similarity 98.4%; Pred. No. 8.4e-21;			
FEATURES	Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
	QY 1 AATCTTCGACGGGGAGGCTGTTTTCCTCCCTTCCTGTTCCGACCGATTCTCAATA 60			
	Db 125 AATCTTCGACGGGGAGGCTGTTTTCCTCCCTTCCTGTTCCGACCGATTCTCAATA 66			
	QY 61 AAAATCATGATTTCATCAAGTTCATTCCTATACATTATCTTTAATAACGATTTTATG 120			
	Db 65 AAAATCATGATTTCATCAAGTTCGTTCTTATACATTATCTTTAATAACGATTTTATG 6			
	QY 121 CTCGG 125			
	Db 5 CTCGG 1			
	RESULT 12			
	AE016997			
	LOCUS			
DEFINITION	AE016997 Chlamydomophila caviae GPIC section 4 of 4 of the complete genome.			
	ACCESSION			
	VERSION			
	KEYWORDS			
	SOURCE			
	ORGANISM			
	REFERENCE			
	AUTHORS			
	TITLE			
	Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the			
JOURNAL	Chlamydiaceae			
	MEDLINE			
	PUBMED			
	REFERENCE			
	AUTHORS			
	TITLE			
	Submitted (29-OCT-2002) The Institute for Genomic Research, 9712			
	Medical Center Dr. Rockville, MD 20850, USA			
	Location/Qualifiers			
	1..267334			
FEATURES	/organism="Chlamydomophila caviae GPIC"			
	/mol_type="genomic DNA"			
	/isolate="GPIC"			
	/db_xref="taxon:227941"			
	source			
	complement (251..421)			
	/locus tag="CCA00790"			
	complement (251..421)			
	/locus tag="CCA00790"			
	/note="identified by Glimmer2; putative"			
FEATURES	/codon_start=1			
	/transl_table=11			
	/product="hypothetical protein"			
	/protein_id="AAP05531.1"			
	/db_xref="GI:29834897"			
	/translation="MEEKGPIKALFVLGEFHHSKNTVECKNMFESRKFTLYTLCKDFF			
	IIVLCFPRSLF"			
	451..1737			
	/locus tag="CCA00791"			
	451..1737			
FEATURES	/locus tag="CCA00791"			
	/note="similar to GB:X69464, PID:398210, and PID:398211;			
	identified by sequence similarity; putative"			
	/codon_start=1			
	/transl_table=11			
	/product="poly(A) polymerase family protein"			
	/protein_id="AAP05532.1"			
	/db_xref="GI:29834898"			
	/translation="MVCDNKTLLFRGELFRKISKAPTIIYSAADHNKIKLDPSPH			
	ALSVKTLKAGHKAVIVGCIJRDLLNTTPKDPDIDISTAKPEIKAVKNCVLLGKR			
FEATURES	FRLAHRFNSQNIENSTFRSGADEDLIITKMLNGTAEDVLRDFTINGFYDPAE			
	ETIITDYGVSDDLQNYLRTIGDPFVRKQDPVRLRLKILARASFTVDPKLEALQ			
	ECYELVKGAPARVFEELIKMLGSGASSEFFKLTIVKYQMLEILFPYMDKAPRLKTLQ			
	EQFSLDVLKDNVSKYNDYRHLQMAAFLPPIVNFVRYKRRQPSLSLTAIFDYI			
	KNFLGQFADSPSSCKKNFIILALVLMQYRLTPLVPTKKIHFNKRLNHRFSEA			
	LSLLEIRSLVYPKLDKILAAVVRHYQALQCKKELPS"			
	1751..3631			
	/gene="lpxB"			
	/locus tag="CCA00792"			
	1751..3631			
FEATURES	/gene="lpxB"			
	/locus tag="CCA00792"			
	/note="similar to GB:X69464, PID:398210, and PID:398211;			
	identified by sequence similarity; putative"			
	/codon_start=1			
	/transl_table=11			
	/product="lipid-A-disaccharide synthase"			
	/protein_id="AAP05533.1"			
	/db_xref="GI:29834899"			
	/translation="MLPLVHVLPVPIGLIANLFTGSFTIOWLISEKBTAVVPKAF			
FEATURES	WLSSTGAVMTAHGFIQSQFPMALLHGANLIVYFNALNIASSYKLSLTITVLIVLT			
	LLVITDFFALAAIYPIYMEWASPNNFFHLPLEPPNIYWHIVGCLGFTFSRRFIQWC			
	YLEMNNHSTLPALFWOAGFVGGLAFIYFIRTDPNVILSYGGLFTFSRRFIQWC			
	SLPLKPHSSTCLISAGSPSGDTLGSDLLENIKELNPNHFCVGGFPMKEGLEPIIR			
	MBEFQSGLEVEFCVAFVSKYKLYKAILKENPETVFCIDFPFHFLIKRLKCG			
	YRGKITHYVCPGIWAPRNKRIKLEKLDLTLILPFEKEIFKDSPLKTIYLGHPLVK			
	TIANFODCNAMKQLEISQPSVALFGSGRPGDIFRNLOVQARAFSSSLAKSHOLLV			
	SSCNPKYDKLLELDKEGCHNNKIVPSKRYQLMRDCCALAKCGTIVLEALNPTP			
	TIVTCLLRFDTLAKITFIKFIYPSLNIITGVSIFPEFIFGGHDFSEVAAAIID			
	ILANPIGSKQKYACQLLKTNTNVTPEKCLQAIYAQNRFYLNKDFIKFHPKSS			
FEATURES	RA"			
	complement (3942..4700)			
	/locus tag="CCA00793"			
	complement (3942..4700)			
	/locus tag="CCA00793"			
	/note="similar to GP:6382503; identified by sequence			
	similarity; putative"			
	/codon_start=1			
	/transl_table=11			
	/product="hypothetical protein"			
FEATURES	/protein_id="AAP05534.1"			
	/db_xref="GI:29834900"			
	/translation="MRVHAMNSVSLVNPSTTNKFMHRFKKCLLSHTAAVAVGIV			
	LILASVALIAYASLLPILCSILFTSALIGLVKYVRYTSLKIDFVDSHERA			
	GYCAMIKSLBIQIDQVRNEQAGDIKVEQLQEVKTLHSHITALQFKEWKQNOQ			
	KEWCAHFDSDRAVELKAVKLYEANKYEQORLOIPELKVAREVRQVILEWVQL			
	RLSRDLOVAANMIQFLQEAEEKVPPKPRSASI"			
	complement (4728..5849)			
	gene			


```

/function="miscellaneous; hypothetical/global homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence-not experimental
/transl_table=11

Query Match      16.1%; Score 43.4; DB 1; Length 92413;
Best Local Similarity 49.8%; Pred. No. 0.96;
Matches 110; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 6 TTCGCCACGGGAGGCTGTTTCTCTCCCTCTGTCGACCGCATTCCTCAATAAATAAT 65
Db 82192 TGGCGAGACGATACCTTTCTTCGTTCTCTTTCAGCTCGGCATATACAGACTAGAT 82133

QY 66 CATTGATTCATCGAAGTTCATCCCTATACCATATCTTTAATAAGATTTATGCTCCG 125
Db 82132 CAGCGAAGACCTCTGGCGTCATTCATCCATTCGCTGCTCGAATTTATTTCTTCAAAC 82073

QY 126 GTTATTCGAATAACCTTAACCTTCACCTCGTAGCAGATCATCGTAGGCTTGGCTATCA 185
Db 82072 GTGGGCAACCGTCTGAAATACTACAGAGTTGCGTCAGCGTCATAGGATAGATATCA 82013

QY 186 ACTCGGCAATCGCAGAAACAGTGTGCGAATACAACTTTTAC 226
Db 82012 ATTCCGCCAAGCGAAGCAGGTTGGCGATAGATTCCTCC 81972

RESULT 15
HGDIIM/c      4884 bp      DNA      linear      BCT 30-JUN-1993
DEFINITION    H.giganteus DIIM gene for methyltransferase.
ACCESSION     X55141.1 S65046
VERSION       X55141.1 GI:48771
KEYWORDS      methyltransferase.
SOURCE        Herpetosiphon aurantiacus
ORGANISM      Herpetosiphon aurantiacus
Bacteria; Chloroflexi; Herpetosiphonales; Herpetosiphonaceae;
Herpetosiphon.
REFERENCE      1 (bases 1 to 4884)
AUTHORS       Kroeger,M.
TITLE         Direct Submission
JOURNAL       Submitted (27-NOV-1990) M. Kroeger, INST MIKRO- UND
              MOLEKULARBIOLOGIE, FRANKFURTER STRASSE 107, D-W-6300 GIESSEN, FRG
              2 (bases 1 to 4884)
REFERENCE      Dusterhoft,A. and Kroeger,M.
AUTHORS       Cloning, sequence and characterization of
              m5C-methyltransferase-encoding gene, hgidiim (GTCGAC), from
              Herpetosiphon giganteus strain Hpa2
              Gene 106 (1), 87-92 (1991)
JOURNAL       92039068
MEDLINE       1937045
PUBMED
COMMENT       See also X55137-X55143.
FEATURES      Location/Qualifiers
              1..4884
               /organism="Herpetosiphon aurantiacus"
               /mol_type="genomic DNA"
               /strain="Hpa2"
               /db_xref="taxon:65"
               884..1300
               /gene="orf 15"
               884..1300
               /gene="orf 15"
               /codon_start=1
               /transl_table=11
               /protein_id="CAA38940.1"
               /db_xref="GI:48772"
               /db_xref="SWISS-PROT:P25279"
               /translation="MIATGAYLWTIREAIGLCRNDVAHEAGTNVQIMRIKGSIDTR
               GSLLSVRAVNFNAEHIAQLFIMLVATEEDGRNLAIWINRETTSAIDFIAVKKD

```

```

NKVSALKLIQOLEALDPSSLDRLLLGYQSLLDNR"
1749..2813
/gene="hgidiim"
CDS
1749..2813
/gene="hgidiim"
/codon_start=1
/transl_table=11
/product="methyltransferase"
/protein_id="CAA38941.1"
/db_xref="GI:48773"
/db_xref="SWISS-PROT:P25265"
/translation="MVGAVIDLFCGVGLTHGLILEGVLVLAGINDPSCKYAEQNN
RTRFEKISVEGDRENLALPNNQHKLIVGCAFCQDFSQYTKSRGTGWQLLTFES
RUIRIEPIIISMEVNEVRNFRGEVFNFIQLEQLGYHSHSVHCPDYGIPQOR
DLVLFAAKQGVKIIIPHTHPENYVRDVIGSLATNYSGHWEGDSMAASLEDI
NURRLOHSPVPGTWADWPELLIAECHKESGESYGVYGRMWDKVAPTITTCNGYV
NGRFHPQDRAISLREAAALQTPRSYQFAPEGQLKKTVSRIQGNVPAVGRIA
KSIKFLGLEHERQVRVRII"
2782..4617
/gene="orf 68"
CDS
2782..4617
/gene="orf 68"
/codon_start=1
/transl_table=11
/protein_id="CAA38942.1"
/db_xref="GI:48774"
/db_xref="SWISS-PROT:P25280"
/translation="MSDSGYELSPSLNTLHSLGYSNILRFYKLVANAWDADATE
VSINRQDEIVIQNGIGMSIEDANTKFLRIGHQKREDSANTI SGRHVMGKGGILA
IFGIANIAEVYCKDGVPHGFIHKGDIERGISDVTLYRSPSPQALSTESGKII
LREIKSSIGNAEKTLRTLARRFTIINNCFVSIINNTPINDRDLYLVKQVFLWL
GGESSKANFETKLKKSPEITNLDGMSGMTKGMIGTVRPSDIPNEQPVHIFYAHG
XMQEDILLIDITAGLYROYIIGIEADFMDSDDDEDIITSNRQRIKQDPSRYKLLE
YKQIDIMVIATWTKLKREYSPKPKKEEVNDSLSKDSNSENENINASSDSSTATE
NASSDSSTATENASSETNDGEVDSNPFDDDIPEFSPPPKOEITAFREKNLVNSN
IPNQMKNILYDIQAAVAYKGTGFKACIVMLGALLEGVLMGTIQTDTLVLEIALPO
VFKPLSDLGPRNPKEADRTVLAQYGTGTFSPQCKEIEIELCVQGTNKLGVDIQTVRN
SIHPGSLVKMKQLARFNHQAGVYIAKLHEIINLVILWNPSP"
BASE COUNT 1500 a 882 c 1055 g 1447 t
ORIGIN

Query Match      15.9%; Score 43; DB 1; Length 4884;
Best Local Similarity 56.8%; Pred. No. 1.2;
Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 29 TCTTCCTCTCTCTCGACCGGATTCCTCAATAAATAATCATTTGATTTCGAGGTCATT 88
Db 3026 TCTTCCTCTCTTGATGCCCAATTCGTAGAATTTAGTATTAGCATCTCTCAATTGACATT 2967

QY 89 CCTATACCATTTATCTTTAATAACGATTTTATGCTCCGGTTTATCGCAATAACCTTACTCC 148
Db 2966 CCTATACCATTTATCTTTGATGACTATTTTCATCTTCTGCTATATTATATAGATACCTCAGTT 2907

QY 149 ACTTCCTGAGCACATGCAT 167
Db 2906 GCATCAGCATCCCAAGCAT 2888

Search completed: November 15, 2003, 03:34:04
Job time : 1199.2 secs

```



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 174.258 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-88

Perfect score: 270

Sequence: 1 ATTCTCCGACGGGGGGG.....CTCGAACTCAATTCATT 270

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270	100.0	270	19	AAV03596
2	270	100.0	270	21	AAV03596
3	253.2	93.8	855	25	AB237755
4	253.2	93.8	858	21	AAV03596
5	253.2	93.8	891	25	AB241353
6	251.6	93.2	858	21	AAV03596
7	121.8	45.1	269	19	AAV03596
8	121.8	45.1	269	21	AAV03596

C 9	42.6	15.8	4951	23	ABL13122
C 10	42.6	15.8	5395	23	ABL15458
C 11	38.2	14.1	6068	24	ABN80204
C 12	37.6	13.9	1230025	20	AAV03596
C 13	37	13.7	7160	20	AAV03596
C 14	36	13.3	4519	19	AAV03596
C 15	36	13.3	11165	21	AAV03596
C 16	35.4	13.1	7529	23	ABL0486
C 17	34.6	12.8	1437	19	AAV04212
C 18	34.6	12.8	1437	24	ABX67809
C 19	34.6	12.8	5829	24	ABX67855
C 20	34.6	12.8	5933	24	ABX39977
C 21	34.6	12.8	6013	24	AAV03596
C 22	34.6	12.8	6013	24	ABX31360
C 23	34.6	12.8	7980	19	AAV04243
C 24	34.6	12.8	10151	24	ABL34433
C 25	34.4	12.7	543	22	ABA61500
C 26	34.4	12.7	543	22	ABA29225
C 27	34.4	12.7	543	22	AAK09799
C 28	34.4	12.7	543	22	AAK35693
C 29	34.4	12.7	543	22	AAV03596
C 30	34.4	12.7	543	23	ABX35412
C 31	34.4	12.7	543	24	ABX09972
C 32	34.2	12.7	1881	24	ABX09528
C 33	34.2	12.7	2365589	24	ABA90521
C 34	34.2	12.7	2944528	24	ABA03041
C 35	34	12.6	15767	24	ABL33206
C 36	34	12.6	15767	24	ABL34552
C 37	33.8	12.5	1327	24	ABL99913
C 38	33.4	12.4	285	21	AAV03596
C 39	33.4	12.4	285	21	AAV03596
C 40	33.2	12.3	13807	18	AAV03596
C 41	33.2	12.3	14602	18	AAV03596
C 42	33.2	12.3	111309	20	AAV03596
C 43	33.2	12.3	910715	20	AAV03596
C 44	33	12.2	5360	24	ABL92297
C 45	33	12.2	5360	24	ABL49366

ALIGNMENTS

RESULT 1

AAV03596
ID AAV03596 standard; DNA; 270 BP.

AC AAV03596;

DT 22-OCT-1998 (first entry)

XX Neisseria meningitidis DNA sequence E24.

XX N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
XX region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
XX meningitis; ss.

OS Neisseria meningitidis.

PN WO9802547-A2.

PD 22-JAN-1998.

PF 11-JUL-1997; 97WO-FR01295.

PR 12-JUL-1996; 96FR-0008768.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (BLAC) MAX PLANCK GSS FORDERUNG WISSENSCHAFTEN.
PA (SMIK) SMITHKLINE BEECHAM.

XX Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI Vinals C;
XX


```

XX PA (CHIR-) CHIRON SPA.
XX PA Fontana MR, Pizza M, Massignani V, Monaci E;
PI PA WPI; 2003-058415/05.
XX DR P-PSDB; ABP76785.
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX DR PT medicament for treating or preventing N. gonorrhoeae infection -
XX PS Claim 6; Page 188; 815pp; English.
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX CC antibodies that specifically bind to the proteins. The composition
XX CC comprising the protein, nucleic acid or antibody is useful for the
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae
XX CC infection, this may be in the form of a vaccine or gene therapy.
XX CC Sequences given in records AB237706-AB242016 represent nucleic acid
XX CC molecules of the invention.
XX SQ Sequence 855 BP; 300 A; 151 C; 178 G; 226 T; 0 other;
Query Match 93.8%; Score 253.2; DB 25; Length 855;
Best Local Similarity 98.5%; Pred. No. 4.6e-65;
Matches 266; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 AATTCCTCCGACGGGAGGCTGTTTCTCCCTCTGTTCCGACCGATTCCTCAATA 60
Db 282 AATTCCTCCGACGGGAGGCTGTTTCTCCCTCTGTTCCGACCGATTCCTCAATA 223
QY 61 AAATCATGATTTCATCGAAGTTCATTCCTATACCATTAATCTTTAATACGATTTTATG 120
Db 222 AAATCATGATTTCATCGAAGTTCATTCCTATGCAATTAATCTTTAATACGATTTTATG 163
QY 121 CTCGGTTTATCGAATACCTTAACCTCCACTTCGATGCAATGATGATGATGATGATG 180
Db 162 CTCGGTTTATCGAATACCTTAACCTCCACTTCGATGCAATGATGATGATGATGATG 103
QY 181 TATCAACTCGGCAATCGGAGGACAGTGTGCAATACAAATCTTTACACCCAAATGTTTGA 240
Db 102 TATCAACTCGGCAATCGGAGGACAGTGTGCAATAC-ATCTTTACACCCAAATGTTTGA 44
QY 241 TTACGGTTGGCTCGAAATCTCAATT 270
Db 43 TTACGGTTGGCTCGAAATCTCAATT 14
RESULT 4
AAA15322/c
ID AAA15322 standard; DNA; 858 BP.
XX AC AAA15322;
XX DT 04-SEP-2000 (first entry)
XX DE DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX OS Neisseria gonorrhoeae.
XX FH Key Location/Qualifiers
XX CDS 1...858
XX FT /*tag= a
XX TT
XX WO200026375-A2.
XX PN 11-MAY-2000.
XX PD
XX PF 28-OCT-1999; 99WO-FR02643.
XX PR
XX

```

```

PR 30-OCT-1998; 98FR-0013693.
XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA (INMR ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PI
XX PI Aujame L, Bouchardon A, Renaud-Mongenie G, Rokbi B, Nassif X;
XX PI Tinsley C, Perrin A;
XX DR WPI; 2000-365622/31.
XX DR P-PSDB; AAY93292.
XX PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX PT or preventative vaccines and for diagnosis -
XX PS Claim 4; Page 134-135; 187pp; French.
XX CC The present sequence encodes a protein that is specific for pathogenic
XX CC strains of Neisseria. The polynucleotides, polypeptides, or their
XX CC antigenic fragments, are used in vaccines to treat or protect against
XX CC Neisseria infections, particularly by N. meningitidis. The
XX CC polynucleotide sequence is also used for recombinant production of
XX CC the polypeptide and to produce attenuated Neisseria strains that
XX CC overexpress it, or express it in a non-toxic mutant form.
XX SQ Sequence 858 BP; 302 A; 151 C; 178 G; 227 T; 0 other;
Query Match 93.8%; Score 253.2; DB 21; Length 858;
Best Local Similarity 98.5%; Pred. No. 4.7e-65;
Matches 266; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 AATTCCTCCGACGGGAGGCTGTTTCTCCCTCTGTTCCGACCGATTCCTCAATA 60
Db 282 AATTCCTCCGACGGGAGGCTGTTTCTCCCTCTGTTCCGACCGATTCCTCAATA 223
QY 61 AAATCATGATTTCATCGAAGTTCATTCCTATACCATTAATCTTTAATACGATTTTATG 120
Db 222 AAATCATGATTTCATCGAAGTTCATTCCTATGCAATTAATCTTTAATACGATTTTATG 163
QY 121 CTCGGTTTATCGAATACCTTAACCTCCACTTCGATGCAATGATGATGATGATGATG 180
Db 162 CTCGGTTTATCGAATACCTTAACCTCCACTTCGATGCAATGATGATGATGATGATG 103
QY 181 TATCAACTCGGCAATCGGAGGACAGTGTGCAATACAAATCTTTACACCCAAATGTTTGA 240
Db 102 TATCAACTCGGCAATCGGAGGACAGTGTGCAATAC-ATCTTTACACCCAAATGTTTGA 44
QY 241 TTACGGTTGGCTCGAAATCTCAATT 270
Db 43 TTACGGTTGGCTCGAAATCTCAATT 14
RESULT 5
AB241353/c
ID AB241353 standard; DNA; 891 BP.
XX AC AB241353;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 7295.
XX KW N. gonorrhoeae infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB02069.
XX PR 12-FEB-2001; 2001GB-0003424.
XX

```



```

XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
XX PI Vinals C;
XX DR WPI; 1998-110594/10.
XX XX
XX PT Genes present in Neisseria meningitidis but not other Neisseria
XX PT species - and related host cells, RNA, anti-sense sequences,
XX PT polypeptide(s) and antibodies, useful for diagnosing Neisseria
XX PT meningitidis infection and in protective vaccines
XX PS
XX PS Example 4; Page 118; 150pp; French.
XX XX
XX CC AAV03575-606 represent sequences that are present in Neisseria
XX CC meningitidis and N. gonorrhoeae but not in N. lactamica, except for the
XX CC genes involved in biosynthesis of the capsule polysaccharide, fcpA or C,
XX CC opc, porA, rotamase, sequence IC1106, IGA protease, pillin, pilC,
XX CC proteins which bind transferrin and opacity proteins. The DNA sequences
XX CC are responsible for the differences in pathogenicity between N.
XX CC meningitidis and N. gonorrhoeae, specifically they include the genes that
XX CC allow N. meningitidis to cross the blood-brain barrier. DNA sequences
XX CC common to N. meningitidis and N. gonorrhoeae, but absent from N.
XX CC lactamica, are responsible for colonisation and penetration of the
XX CC mucosa. The DNA sequences can be used to produce probes and primers, and
XX CC antibodies produced against the encoded proteins are used in standard
XX CC hybridisation/immunoassay processes for diagnosis of N. meningitidis
XX CC infection, particularly meningitis.
XX SQ
XX Query Match 45.1%; Score 121.8; DB 19; Length 269;
XX Best Local Similarity 98.4%; Pred. No. 2.4e-26;
XX Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AATCTTCGCGACGGGGAGGCTTGTTTCTTCCTCTCTGTCGACCGATTCTCAATA 60
Dd 125 AATCTTCGCGACGGGGAGGCTTGTTTCTTCCTCTCTGTCGACCGATTCTCAATA 66
QY 61 AAATCATGATTTCATCGAGTTCATTCCTATACCATATCTTTAATACGATTTTATG 120
Dd 65 AAATCATGATTTCATCGAGTTCATTCCTATACCATATCTTTAATACGATTTTATG 6
QY 121 CTCGG 125
Dd 5 CTCGG 1

RESULT 8
AAAL5389/c
ID AAA15389 standard; DNA; 269 BP.
XX AC
XX AC AAAL5389;
XX XX
XX DT 04-SEP-2000 (first entry)
XX DE
XX DE Genomic fragment of Neisseria meningitidis Z2491.
XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX OS Neisseria meningitidis.
XX XX
XX PN WO200026375-A2.
XX XX
XX PD 11-MAY-2000.
XX XX
XX PF 28-OCT-1999; 99WO-FR02643.
XX XX
XX PR 30-OCT-1998; 98FR-0013693.
XX XX

(PA (INRM ) PASTEUR MERIEUX SERUMS & VACCINS SA.
(PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI Tinsley C, Ferrin A;
XX DR WPI; 2000-365622/31.
XX XX
XX PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX PT or preventative vaccines and for diagnosis
XX PS
XX PS Claim 1; Page 40; 187pp; French.
XX XX
XX CC The present sequence represents a genomic fragment of Neisseria
XX CC meningitidis. The specification describes proteins that are specific
XX CC for pathogenic strains of Neisseria. The polynucleotides, polypeptides,
XX CC or their antigenic fragments, are used in vaccines to treat or protect
XX CC against Neisseria infections, particularly by N. meningitidis. The
XX CC polynucleotide sequences are also used for recombinant production of
XX CC the polypeptide and to produce attenuated Neisseria strains that
XX CC overexpress it, or express it in a non-toxic mutant form.
XX CC note: the present sequence is included in Claim 1, but it is those
XX CC sequences that do not include the present sequence that are actually
XX CC claimed.
XX SQ
XX Query Match 45.1%; Score 121.8; DB 21; Length 269;
XX Best Local Similarity 98.4%; Pred. No. 2.4e-26;
XX Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AATCTTCGCGACGGGGAGGCTTGTTTCTTCCTCTCTGTCGACCGATTCTCAATA 60
Dd 125 AATCTTCGCGACGGGGAGGCTTGTTTCTTCCTCTCTGTCGACCGATTCTCAATA 66
QY 61 AAATCATGATTTCATCGAGTTCATTCCTATACCATATCTTTAATACGATTTTATG 120
Dd 65 AAATCATGATTTCATCGAGTTCATTCCTATACCATATCTTTAATACGATTTTATG 6
QY 121 CTCGG 125
Dd 5 CTCGG 1

RESULT 9
ABLJ3122/c
ID ABLJ3122 standard; cDNA; 4951 BP.
XX AC
XX AC ABLJ3122;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 33848.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX XX
XX PN WO200171042-A2.
XX XX
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX XX
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX XX
XX PA (PEXE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.

```

DR P-PSDB; ABB69019.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell

PT interactions -

XX

XX Claim 1; SEQ ID NO 33848; 21pp + Sequence Listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 4951 BP; 1376 A; 1121 C; 1065 G; 1389 T; 0 other;

Query Match 15.8%; Score 42.6; DB 23; Length 4951;

Best Local Similarity 61.1%; Pred. No. 0.016; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 44;

QY 30 CTTCCCTTCTGTCGACCGATTCTCAATAAAATCATGATTCATCGAAGTTCATTC 89

DB 4057 CATGCACCTCTGTAACCGTATACCTACAAAATATATCATATATTAACAAATTTATCT 3998

QY 90 CTATACCATTTATCTTTATATACGATTTTATGCTCCGTTTATCGAATAACCTA 142

DB 3997 CTTTACCATTATACTAAACGATGTGATGCTCGGATACACTGCATCATT 3945

RESULT 10

ABL15458/c

ID ABL15458 standard; cDNA; 5395 BP.

XX

XX ABL15458;

XX

XX 26-MAR-2002 (first entry)

XX

XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 40856.

XX

XX *Drosophila*; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX

XX *Drosophila melanogaster*.

XX

XX WO200171042-A2.

XX

XX 27-SEP-2001.

XX

XX 23-MAR-2001; 2001WO-US09231.

XX

XX 23-MAR-2000; 2000US-191637P.

XX

XX 11-JUL-2000; 2000US-0614150.

XX

XX (PEXE) PE CORP NY.

XX

XX Venter JC, Adams M, Li PWD, Myers EW;

XX

XX WPI; 2001-656860/75.

XX

XX P-PSDB; ABB71355.

XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell

PT interactions -

XX

XX Claim 1; SEQ ID NO 40856; 21pp + Sequence Listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

CC

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 5395 BP; 1528 A; 1109 C; 1141 G; 1617 T; 0 other;

Query Match 15.8%; Score 42.6; DB 23; Length 5395;

Best Local Similarity 61.1%; Pred. No. 0.016; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 44;

QY 30 CTTCCCTTCTGTCGACCGATTCTCAATAAAATCATGATTCATCGAAGTTCATTC 89

DB 536 CATGCACCTCTGTAACCGTATACCTACAAAATATATCATATATTAACAAATTTATCT 477

QY 90 CTATACCATTTATCTTTATATACGATTTTATGCTCCGTTTATCGAATAACCTA 142

DB 476 CTTTACCATTATACTAAACGATGTGATGCTCGGATACACTGCATCATT 424

RESULT 11

ABN80204/c

ID ABN80204 standard; DNA; 6068 BP.

XX

XX ABN80204;

XX

XX 15-JUL-2002 (first entry)

XX

XX Human chemically modified disease associated gene SEQ ID NO 221.

XX

XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;

XX heart disease; epilepsy; histone deacetylation; muscular dystrophy;

XX dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;

XX antidiabetic; cytostatic; anticonvulsant; ds.

XX

XX Homo sapiens.

XX

XX Synthetic.

XX

XX WO200200927-A2.

XX

XX 03-JAN-2002.

XX

XX 02-JUL-2001; 2001WO-EP07536.

XX

XX 30-JUN-2000; 2000DE-1032529.

XX

XX 01-SEP-2000; 2000DE-1043826.

XX

XX (EPIG-) EPIGENOMICS AG.

XX

XX Olek A, Piepenbrock C, Berlin K;

XX

XX WPI; 2002-130908/17.

XX

XX Novel nucleic acid useful for diagnosis and therapy of diseases

XX associated with development genes such as diabetes, comprises a

XX sequence of a segment of chemically pretreated DNA of genes associated

XX with development -

XX

XX Claim 1; SEQ ID NO 221; 27pp; English.

XX

XX The invention relates to a nucleic acid (i) comprising a sequence at

CC least 18 bases in length of a segment of chemically pretreated DNA (ii)

CC of genes associated with development selected from 87 genes listed in

CC the specification such as ACCPN, APFN, or AP1 and comprising one of 350

CC sequences (ABN7994-ABN8033) or their complements. The invention is

CC useful for the diagnosis or therapy of diseases associated with

CC development genes, in particular disease related to homeobox containing

CC genes (Hox), like diabetes, cancer, apoptosis related diseases, syndromes
 CC associated with congenital heart disease, epilepsy, diseases related to
 CC histone deacetylation, Currarino syndrome, diseases related with the
 CC development of the brain and limb girdle muscular dystrophy and dwarfism.
 CC Oligomers specific to each of the genes are useful for detecting the
 CC methylation state of all CpG dinucleotides within the 350 sequences or
 CC (II) and their complementary sequences, as primer oligonucleotides for
 CC the amplification of the 350 sequences, (II) and/or their complements and
 CC as oligomer probes for detecting the cytosine methylation state and/or
 CC single nucleotide polymorphisms (SNPs).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.

CC Sequence 6068 BP; 1819 A; 81 C; 1390 G; 2778 T; 0 other;
 CC
 CC Query Match 14.1%; Score 38.2; DB 24; Length 6068;
 CC Best Local Similarity 59.8%; Pred. No. 0.34;
 CC Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 CC
 CC QY 56 AATAAATAATCATTCATTCGGAAGTTCTCTATACCATTTATTAATACGATT 115
 CC Db 4187 AAAAAAATAATTAACATAATAAATCTTCTACATTTATTTATATACCATTA 4128
 CC
 CC QY 116 TTATGCTCCGGTTTATCGAATAACCTAATTCCTCCACTCCGTCAGCACA 162
 CC Db 4127 TTATCCTACCGCCTTATCAAAAACACCACTAAATCCTCACTAACACA 4081

CC RESULT 12
 CC AAX91990/c
 CC ID AAX91990 standard; DNA; 1230025 BP.
 CC AC AAX91990;
 CC XX
 CC DT 13-SEP-1999 (first entry)
 CC XX
 CC DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
 CC KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 CC KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 CC KW vaccine; neutralising epitope; ss.
 CC XX
 CC OS Chlamydia pneumoniae.
 CC XX
 CC PN WO9927105-A2.
 CC PD 03-JUN-1999.
 CC PF 20-NOV-1998; 98WO-IB01890.
 CC PR 04-NOV-1998; 98US-0107078.
 CC PR 21-NOV-1997; 97FR-0014673.
 CC XX
 CC PA (GEST) GENSET.
 CC XX
 CC PI Grifais R;
 CC XX
 CC DR WPI; 1999-357842/30.
 CC DR
 CC PT Genome sequence of Chlamydia pneumoniae
 CC XX
 CC PS Claim 1; Page 291-611; 1912pp; English.
 CC XX

CC The present sequence represents the complete genome of Chlamydia
 CC pneumoniae, and encodes proteins AAY34584-Y33879. C. pneumoniae causes
 CC respiratory disease such as pneumonia and bronchitis and is thought
 CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
 CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
 CC encoded by the open reading frames of the C. pneumoniae genome (see
 CC AAY34584-Y33879) can be used in immunogenic compositions as vaccines.
 CC Vectors containing C. pneumoniae nucleotides sequences can also be
 CC used as immunogenic compositions, especially where the vector directs

CC the expression of a neutralising epitope of C. pneumoniae.

CC XX Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;
 CC SQ

CC Query Match 13.9%; Score 37.6; DB 20; Length 1230025;
 CC Best Local Similarity 54.3%; Pred. No. 2.5;
 CC Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

CC QY 70 GATTTCATCGAAGTTTCATTCCTATACCATTTCTTAATAAGATTTTATGTCCTCGGTTT 129
 CC Db 929463 GATGCTTTCGGCTCTGAAGCCACCACTATCTCTGATAATGATTCGCGCTTGTCCTCC 929404
 CC QY 130 ATCGAATAACCTAACTTCCACTTCCGTAGCATGTCATGCTAGGCATTCGCTATCAACTC 189
 CC Db 929403 TCTTAAAGTTTCGATTCTTATTCATCGCGCCAGCATCTAAGGAATTCCTCATCACTTC 929344
 CC QY 190 GCGAATCGCAGNACAGTGT 209
 CC Db 929343 TTTTAAACACAGAAACGGAGT 929324

CC RESULT 13
 CC AAX86191
 CC ID AAX86191 standard; DNA; 7160 BP.
 CC XX
 CC AC AAX86191;
 CC XX
 CC DT 22-SEP-1999 (first entry)
 CC XX
 CC DE Nucleotide sequence of the TONNEAU1 gene.
 CC XX

CC KW TONNEAU1; cell division; cell elongation; eukaryotic cell; plant cell;
 CC KW floral structure; male sterility; anther dehiscence; petal removal;
 CC KW apetalous seed; female sterility; ornamental plant; ss.

CC OS Arabidopsis thaliana.

CC PN WO935269-A2.

CC XX 15-JUL-1999.

CC PD 08-JAN-1999; 99WO-FR00021.

CC PF 09-JAN-1998; 98FR-0000162.

CC PR (INRG) INRA INST NAT RECH AGRONOMIQUE.

CC PA Bouchez D, Nacry P;

CC PI WPI; 1999-430396/36.

CC DR P-PSDB; AAY23987, AAY23988.

CC PT Controlling division and elongation of eukaryotic plant cells

CC XX Example 2; Page 42-45; 48pp; French.

CC The present sequence represents the TONNEAU1 gene of Arabidopsis
 CC thaliana. The specification describes a method to control division
 CC and/or elongation of eukaryotic cells, and in particular plant
 CC cells, by modifying the expression and/or activity of a TONNEAU1
 CC protein in the cells. The method is useful for the control of division
 CC and/or elongation of plant cells. The TONNEAU1 protein coding
 CC sequence can be inhibited by an antisense sequence or by mutagenesis.
 CC The expression of TONNEAU1-type genes can be tissue specific.
 CC Alternatively, TONNEAU1 protein inhibitors are used to control its
 CC activity. In particular, the method is used to alter floral structure
 CC to induce male sterility (by reducing size and blocking dehiscence of
 CC anthers), removal of petals (by producing an apetalous seed) or
 CC inducing female sterility in ornamental species at the end of the
 CC flowering period.

CC SQ Sequence 7160 BP; 2086 A; 1287 C; 1349 G; 2438 T; 0 other;

Query Match 13.7%; Score 37; DB 20; Length 7160;

Best Local Similarity 58.7%; Pred. No. 0.8; Mismatches 45; Indels 0; Gaps 0;

Matches 64; Conservative 0;

QY 28 TTCTTCCTCTCTGTTCCGACGATTTCTCAAAATAAAATCATTTGATTTTCATCGAAGTTTCAT 87

DB 427 TTCTTCATTCGGAATTCATCCCATTTGACAAACCAACGCGTTGGATCACTCAATGCTCAT 486

QY 88 TCCTATACCATATCTTTAATAACGATTTATGCTCCGGTTTATCGAAT 136

DB 487 TCGTAAACACGATATCCCTCATTAGCTTTTCTTTCCCATTTTGGTTGAAT 535

RESULT 14

AAV31192 standard; DNA; 4519 BP.

XX AC AAV31192;

DT 01-OCT-1998 (first entry)

DE E. coli J96 pathogenicity island contig #6.

XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;

KW PAI V; pheV; vaccine; protective immune response; ds.

XX Escherichia coli.

XX WO982575-A2.

XX 28-MAY-1998.

XX 21-NOV-1997; 97WO-US21347.

XX 14-OCT-1997; 97US-0061953.

XX 22-NOV-1996; 96US-0031626.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (UYWI-) UNIV WISCONSIN.

XX Choi GH, Dillon PJ, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to

PT develop products for the detection of pathogenic E. coli and to

PT elicit an immune response to pathogenic E. coli

XX Claim 21; Page 94-96; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing

CC pathogenicity island (PAI) sequences, and represents a nucleic acid

CC molecule of the invention. PAIs are large fragments of DNA which comprise

CC pathogenicity determinants. The sequences of the invention are taken from

CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)

CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at

CC approximately 94 min (at pheR) on the E. coli chromosome and is

CC approximately 160 kb in size. Antibodies specific to the proteins encoded

CC by the PAI open reading frames of the invention can be used in kits to

CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit

CC a protective immune response in an animal to the uropathogenic E. coli

XX strain J96.

XX Sequence 4519 BP; 1141 A; 1101 C; 918 G; 1352 T; 7 other;

QY Query Match 13.3%; Score 36; DB 19; Length 4519;

Best Local Similarity 56.4%; Pred. No. 1.4;

Matches 88; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 47 CCGATTCTCAATAAAATGATTTTCATCGAGTTTCATTCCTTATACCATTCATTTA 106

DB 95 CCAATAGTCAACACCTCTCTTCAAAATTCATCGGTTGTCATACCTATCCATCACT--- 151

QY 107 ATACGATTTTATGCTCCGGTTTATCGAATAACCTAACCTTCCACTTCCGTAGCACATGCA 166

DB 152 CTCAGATAAGAGATTTTCTTCTCTAAATAAATCAACTTCGACATTATCAGCATAGGCA 211

QY 167 TCGTAGGATTCGGTATCACTCAATCGGCAATCGCAGGA 202

DB 212 TCATGAGCATTTTAAATAAATCACTCAATCAAGGCAGTA 247

RESULT 15

AAAI5186/c

ID AAAI5186 standard; DNA; 11165 BP.

XX AC AAAI5186;

DT 04-SEP-2000 (first entry)

DE DNA encoding Escherichia coli virulence proteins.

XX Virulence protein; tatA; tatB; tatC; tatE; mdoG; creC; recG; yggN;

KW eckA; iroD; iroC; iroE; mtg2; msl1; vaccine; infection;

KW Gram negative bacterium; ss.

XX Escherichia coli.

XX Location/Qualifiers

Key 2..1099

CDS /tag= a

FT /product= "virulence protein"

FT /note= "encodes AAY93237; no termination codon given"

FT 1102..1488

FT /tag= b

FT /product= "virulence protein"

FT /note= "encodes AAY93238"

FT 1573..1896

FT /tag= c

FT /product= "virulence protein"

FT /note= "encodes AAY93239"

FT 1939..2196

FT /tag= d

FT /product= "virulence protein"

FT /note= "encodes AAY93240; no termination codon given"

FT 2198..2533

FT /tag= e

FT /product= "virulence protein"

FT /note= "encodes AAY93241"

FT 2613..3041

FT /tag= f

FT /product= "virulence protein"

FT /note= "encodes AAY93242; no termination codon given"

FT 3054..3410

FT /tag= g

FT /product= "virulence protein"

FT /note= "encodes AAY93243"

FT 3460..3705

FT /tag= h

FT /product= "virulence protein"

FT /note= "encodes AAY93244"

FT 3791..4837

FT /tag= i

FT /product= "virulence protein"

FT /note= "encodes AAY93245"

FT 4878..7802

FT /tag= j

FT /product= "virulence protein"

FT /note= "encodes AAY93246"

FT 7816..9483

FT /tag= k

FT /product= "virulence protein"

FT /note= "encodes AAY93247"

FT 9836..10084

FT /tag= l

FT /product= "virulence protein"


```

FT      CDS      /note= "encodes AAY93248"
FT      10134..10430
FT      /*tag= m
FT      /product= "virulence protein"
FT      /note= "encodes AAY93249"
FT      10459..10779
FT      /*tag= n
FT      /product= "virulence protein"
FT      /note= "encodes AAY93250"
XX
XX PN WO200028038-A2.
XX
XX PD 18-MAY-2000.
XX
XX PF 09-NOV-1999; 99WO-GB03721.
XX
XX PR 09-NOV-1998; 98GB-0024569.
XX PR 09-NOV-1998; 98GB-0024570.
XX PR 17-DEC-1998; 98GB-0027814.
XX PR 17-DEC-1998; 98GB-0027815.
XX PR 17-DEC-1998; 98GB-0027816.
XX PR 17-DEC-1998; 98GB-0027818.
XX PR 13-JAN-1999; 99GB-0000708.
XX PR 13-JAN-1999; 99GB-0000710.
XX PR 13-JAN-1999; 99GB-0000711.
XX PR 28-JAN-1999; 99GB-0001915.
XX
XX PA (MICR-) MICROSCIENCE LTD.
XX
XX PI Crooke HR, Clarke EE, Everest PH, Dougan G, Holden DW, Shea JE;
XX PI Feldman RG;
XX
XX DR WPI; 2000-376550/32.
XX DR P-PSDB; AAY93237, AAY93238, AAY93239, AAY93240, AAY93241, AAY93242,
XX DR AAY93243.
XX
XX PT Peptide encoded by an operon including genes from Escherichia coli for
XX PT screening potential drugs, detecting virulence and treating conditions
XX PT associated with infection by a Gram negative bacterium -
XX
XX PS Disclosure; Page 83-101; 122pp; English.
XX
XX CC The present sequence encodes Escherichia coli virulence proteins
XX CC The specification describes virulence proteins which are encoded
XX CC by an operon including tatA, tatB, tatC, tatE, mdoG, creC, recG, yggN,
XX CC eckl, iroD, iroC, iroE, mtd2 or msl-16 genes obtained from Escherichia
XX CC coli K1. The virulence proteins and polynucleotides, and their vaccines
XX CC are useful for screening potential drugs, for the detection of virulence,
XX CC and for treating or preventing conditions associated with infection by
XX CC a Gram negative bacterium particularly Escherichia coli.
XX
XX SQ Sequence 11165 BP; 3592 A; 2087 C; 2458 G; 3028 T; 0 other;
XX
XX Query Match 13.3%; Score 36; DB 21; Length 11165;
XX Best Local Similarity 56.4%; Pred No. 1.8;
XX Matches 88; Conservative 0; Mismatches 65; Indels 3; Gaps 1;
XX
XX QY 47 CCGATTCTCAATAAATATCATTTGATTCATCGAAGTTTCATTCCTATACCATTTATCTTTA 106
XX |||||
XX 5062 CCAATAGTCACACCACCTCTCTCAATTCATCGGTTGTCATACCTATCCATCATCT--- 5006
XX
XX QY 107 ATACGATTTTATGCTCGGTTTATCGAATTAACCTAATTCCTTCCTAGCAGCATGCA 166
XX |||||
XX Db 5005 CTCAGATAAAGAAGATTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4946
XX
XX QY 167 TCGTAGGCAATTCGGTATCATCAATCGGCAATCGCAGGA 202
XX |||||
XX Db 4945 TCATGAGCATTTTAAATAACTCACTCAAGGCAGTA 4910
XX

```

Search completed: November 15, 2003, 00:35:37
 Job time : 178.258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1801.19 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-88
Perfect score: 270
Sequence: 1 AATCTCCGACGGGAGG.....CTCGAACTCAATTCAATT 270

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2781392 seqs, 12152239056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthm:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42.6	15.8	521	10 AW942932	AW942932 LD29090.3
C 2	41.2	15.3	548	9 AV95775	AV95775 AV95775
C 3	40	14.8	722	13 BU437178	BU437178 604146050
C 4	40	14.8	753	9 AJ450902	AJ450902 AJ450902

C 5	40	14.8	856	13 BU421944	BU421944 603955426
C 6	39.4	14.6	637	28 A5224804	A5224804 2365BE08
C 7	39	14.4	437	9 AJ456670	AJ456670 AJ456670
C 8	38.2	14.1	946	13 BQ735322	BQ735322 AGENCOURT
C 9	37.4	13.9	527	12 BM256945	BM256945 520410 MA
C 10	37.2	13.8	1201	13 BX443799	BX443799 BX443799
C 11	37	13.7	280	29 BX293327	BX293327 Arabidops
C 12	37	13.7	1030	29 CC254763	CC254763 CH261-136
C 13	36.8	13.6	688	28 BH934691	BH934691 odes8d03
C 14	36.8	13.6	801	10 BE225447	BE225447 MD0355 Me
C 15	36.8	13.6	809	29 CC122282	CC122282 NDL.20H11
C 16	36.8	13.6	890	28 AZ551079	AZ551079 ENTVD10TR
C 17	36.8	13.6	921	28 BH151796	BH151796 ENTW75TR
C 18	36.6	13.6	492	28 AA550645	AA550645 1834m3 gm
C 19	36.6	13.6	500	9 AU086225	AU086225 AU086225
C 20	36.2	13.4	437	28 A2173796	A2173796 SP.0125_A
C 21	36.2	13.4	446	29 B2417183	B2417183 1f76all1g
C 22	36.2	13.4	708	29 B2653531	B2653531 OGAMF80TM
C 23	36.2	13.4	759	13 BQ704392	BQ704392 Bn01_05K2
C 24	36.2	13.4	777	29 B2391784	B2391784 EINAR96TF
C 25	36.2	13.4	815	29 B2637852	B2637852 OGAMW90TM
C 26	36.2	13.4	1201	13 BX385455	BX385455 BX385455
C 27	36	13.3	441	9 AI485891	AI485891 EST244212
C 28	36	13.3	542	13 BQ487518	BQ487518 WHE2164_C
C 29	35.8	13.3	863	28 BH574782	BH574782 BOG1H33TR
C 30	35.8	13.3	1201	9 AL538100	AL538100 AL538100
C 31	35.8	13.3	1201	13 BX423406	BX423406 BX423406
C 32	35.8	13.3	1293	29 CC293132	CC293132 CH261-190
C 33	35.6	13.2	499	12 BJ108544	BJ108544 BJ108544
C 34	35.6	13.2	517	13 BU237942	BU237942 Hy_ad_04A
C 35	35.6	13.2	1104	13 BX436473	BX436473 BX436473
C 36	35.6	13.2	1201	9 AL526873	AL526873 AL526873
C 37	35.2	13.0	397	13 BX101702	BX101702 BX101702
C 38	35.2	13.0	535	28 AO512978	AO512978 HS_5139_A
C 39	35.2	13.0	590	29 B2421054	B2421054 1f76all1b
C 40	35.2	13.0	634	28 AZ523475	AZ523475 219BDH01
C 41	35.2	13.0	833	29 CNS035LS	AL240853 Tetraodon
C 42	35.2	13.0	1584	29 CC237385	CC237385 CH261-191
C 43	35	13.0	400	11 AY067665	AY067665 Schmidtea
C 44	35	13.0	770	29 CC131165	CC131165 NDL_94G2
C 45	35	13.0	864	28 A2686202	A2686202 ENTWD27TF

ALIGNMENTS

RESULT 1
AW942932
LOCUS 521 bp mRNA linear EST 23-APR-2001
DEFINITION LD29090.3prime LD Drosophila melanogaster embryo pot2 Drosophila melanogaster cDNA clone LD29090 3, mRNA sequence.
ACCESSION AW942932
VERSION AW942932.1 GI:8120637
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 521)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMI Drosophila EST Project.
JOURNAL Unpublished
COMMENT Other_ESTs: LD29090.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd
Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was

polyadenylated. The resulting Poly-T sequence has been removed. hit
 Genomic AE003678: Drosophila melanogaster genomic scaffold
 142000013386035 section 3 of 105, complete sequence.: 02/25/2001
 Plate: LD.290 row: H column: 6
 High quality sequence stop: 382.

FEATURES

Location/Qualifiers
 1..521
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clones="LD29030"
 /sex="male and female"
 /dev_stage="0 to 24 hours mixed stage embryonic"
 /lab_host="XLI Blue"
 /clone_lib="LD Drosophila melanogaster embryo pOT2"
 /note="Organ: embryo; Vector: pOT2; Site 1: EcoRI; Site 2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pOT2."

BASE COUNT 147 a 129 c 100 g 145 t

ORIGIN

Query Match 15.8%; Score 42.6; DB 10; Length 521;
 Best Local Similarity 61.1%; Pred. No. 0.65;
 Matches 69; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 30 CTTCCCTCTCTTCGACCGATTCTCAATAAATAATCATTGATTTCATCGAAGTTTCATTC 89
 |||||
 DB 154 CATGCACCTCTGAACCGTATACCTACAAAATATATCATATATTTAAACAAATTTATCT 213
 |||||
 QY 90 CTATACCATTTATCTTTATACGATTTTATCTCGGTTTATCGAATAACCTA 142
 |||||
 DB 214 CTTTACCATTTATCTTTATACGATTTTATCTCGGTTTATCGAATAACCTA 266
 |||||

RESULT 2

AV995775/c

LOCUS AV995775 548 bp mRNA linear EST 15-MAR-2002
 DEFINITION intestinalis cDNA clone citb42e21 5', mRNA sequence.

ACCESSION AV995775

VERSION AV995775.1

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

REFERENCE 1 (bases 1 to 548)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

Location/Qualifiers

1..548
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="citb42e21"
 /tissue_type="whole animal"
 /dev_stage="tailbud embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, tailbud
 embryo"

BASE COUNT 167 a 95 c 138 g 148 t

ORIGIN

Query Match 15.3%; Score 41.2; DB 9; Length 548;
 Best Local Similarity 57.9%; Pred. No. 1.6;
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 64 ATCATTGATTTTCATCGAAGTTTCATTCCTATACCATTTATCTTTAATAACGATTTTATGTC 123
 |||||
 DB 314 ATATTTCATTTCTTCTGCTTTAATTCACACCATTTATCCCGAATTCGATTTTCAAG 255
 |||||
 QY 124 CGGTTTTCGAATACCTTAACCTTCACCTTCGTCAGACATCATCGTAGGATTCGGTAT 183
 |||||
 DB 254 CCCAAATCGACCAATTTGACTTGAATCGACGTTGCACCGCGTCCAGTGCATTTCTCAGT 195
 |||||
 QY 184 CAACCTC 189
 |||||
 DB 194 TAGTTTC 189
 |||||

RESULT 3

BU437178/c

LOCUS BU437178

DEFINITION

ACCESSION BU437178

VERSION BU437178.1

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 722)

AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Pong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22355534

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..722
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer and broiler"
 /db_xref="taxon:9031"
 /clone="CHEST986c14"
 /sex="Male and female"
 /tissue_type="muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQRBN11"
 /note="Vector: pBluescript II KS(+); Site 1: EcoRI;
 Site 2: NotI; This normalized library was constructed from
 1 million independent clones. cDNA synthesis was initiated
 using an oligo(dT) primer, using methylated C in the first
 strand synthesis reaction. Following this first strand
 reaction, double-stranded cDNA was blunted, ligated to
 NotI adapters, digested with EcoRI, size-selected, and
 cloned into the NotI and EcoRI compatible sites of a
 custom modified MCS of the pBluescript (KS+) vector. The
 library was normalized in 2 rounds using conditions
 adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
 Bonaldo et al., Genome Research 6 (1996): 791, except that
 a significantly longer reannealing hybridization was
 used."

BASE COUNT 224 a 139 c 166 g 192 t 1 others

ORIGIN

Query Match 14.8%; Score 40; DB 13; Length 722;
 Best Local Similarity 54.9%; Pred. No. 3.4;

Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 62 AAATCATGATTTCATCGAGTTCATCTCTATACCATATCTTTAATAACGATTTATGC 121
 Db 171 ATAACAGGAACATCAGCGACCTTGATTCATGCCATTGCTCTTACTTCTATTGTGCA 112
 QY 122 TCCGGTTTATCGAATAACCTAACTCCACTTCCGTAGCACATGCGTAGGCATTGCGCT 181
 Db 111 AACCCGTAATTCAGTCTTAATGTCATAGTAGTGGCACTGGCATCCAAAGAGTTTCT 52
 QY 182 ATCAACTCGGCAATCGCAGGAACA 205
 Db 51 ATCAGCTCCTTCACCACGCTGACA 28

RESULT 4
 AJ450902/c
 LOCUS AJ450902 753 bp mRNA linear EST 19-APR-2002
 DEFINITION AJ450902 riken1 Gallus gallus cdna clone 2622r1, mRNA sequence.
 ACCESSION AJ450902
 VERSION AJ450902.1 GI:20218123
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 753)
 AUTHORS Buerstedde, J.M.
 TITLE Gallus gallus bursal lymphocyte EST
 JOURNAL Unpublished
 COMMENT Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institut
 Martinstr. 52, 20251 Hamburg, Germany
 Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
 FEATURES
 Location/Qualifiers
 source
 1..753
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 /clone="26p22r1"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /clone_lib="riken1"
 /note="CB inbred strain"

BASE COUNT 226 a 151 c 179 g 194 t 3 others

ORIGIN
 Query Match 14.8%; Score 40; DB 9; Length 753;
 Best Local Similarity 54.9%; Pred. No. 3.4;
 Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 62 AAATCATGATTTCATCGAGTTCATCTCTATACCATATCTTTAATAACGATTTATGC 121
 Db 260 ATAACAGGAACATCAGCGACCTTGATTCATGCCATTGCTCTTACTTCTATTGTGCA 201
 QY 122 TCCGGTTTATCGAATAACCTAACTCCACTTCCGTAGCACATGCGTAGGCATTGCGCT 181
 Db 200 AACCCGTAATTCAGTCTTAATGTCATAGTAGTGGCACTGGCATCCAAAGAGTTTCT 141
 QY 182 ATCAACTCGGCAATCGCAGGAACA 205
 Db 140 ATCAGCTCCTTCACCACGCTGACA 117

RESULT 5
 BU421944/c
 LOCUS BU421944 866 bp mRNA linear EST 29-NOV-2002
 DEFINITION 603955426F1 CSEQRBN09 Gallus gallus cdna clone CHES1922c8 5', mRNA
 sequence.
 ACCESSION BU421944
 VERSION BU421944.1 GI:25914615

KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 866)
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.R., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..866
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="layer and broiler"
 /db_xref="taxon:9031"
 /clone="ChEST922c8"
 /sex="Male and female"
 /tissue_type="Chondrocytes isolated from growth plate
 cartilage"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQRBN09"

Site 2: Not1; this normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 261 a 176 c 201 g 228 t

ORIGIN
 Query Match 14.8%; Score 40; DB 13; Length 866;
 Best Local Similarity 54.9%; Pred. No. 3.4;
 Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 62 AAATCATGATTTCATCGAGTTCATCTCTATACCATATCTTTAATAACGATTTATGC 121
 Db 261 ATAACAGGAACATCAGCGACCTTGATTCATGCCATTGCTCTTACTTCTATTGTGCA 202
 QY 122 TCCGGTTTATCGAATAACCTAACTCCACTTCCGTAGCACATGCGTAGGCATTGCGCT 181
 Db 201 AACCCGTAATTCAGTCTTAATGTCATAGTAGTGGCACTGGCATCCAAAGAGTTTCT 142
 QY 182 ATCAACTCGGCAATCGCAGGAACA 205
 Db 141 ATCAGCTCCTTCACCACGCTGACA 118

RESULT 6
 AZ524804
 LOCUS AZ524804 637 bp DNA linear GSS 07-MAY-2001
 DEFINITION 236PBE08 Pb MEN #21 Plasmodium berghei genomic 3', genomic survey
 sequence.

ACCESSION AZ524804
 VERSION AZ524804.1 GI:13965029
 LOCUS GSS.
 DEFINITION Plasmodium berghei
 ACCESSION Plasmodium berghei
 VERSION Plasmodium berghei
 KEYWORDS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 SOURCE 1 (bases 1 to 637)
 ORGANISM Carlon, J.M.-R. and Dame, J.B.
 REFERENCE The Plasmodium vivax and P. berghei gene sequence tag projects
 AUTHORS Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
 TITLE Contact: Dame JB
 JOURNAL Dept. of Pathobiology, College of Veterinary Medicine
 COMMENT University of Florida
 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: damej@mail.vetmed.ufl.edu
 Seq primer: M13(-20) forward
 Class: shotgun.
 FEATURES
 Location/Qualifiers
 1..637
 /organism="Plasmodium berghei"
 /mol_type="genomic DNA"
 /strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
 /db_xref="taxon:5821"
 /dev_stages="asexual blood forms"
 /lab_host="Mus musculus"
 /clone_libs="Pb MEN #21"
 /note="Vector: pBluescript SK(+), phagemid excised from lambda ZAP; Site 1: EcoRV; Site 2: EcoRV; Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from contaminating host DNA by Hoechst Dye 33258-CsCl ultracentrifugation and precipitated. Purified DNA was digested with Mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."
 BASE COUNT 190 a 97 c 51 g 299 t
 ORIGIN
 Query Match 14.6%; Score 39.4; DB 28; Length 637;
 Best Local Similarity 47.4%; Pred. No. 4.8;
 Matches 118; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 18 AGGCTGTTTCTTCCCTTCGTGTCGACCGATCTCAATAAATCAATGATTTTCAT 77
 Db 316 AGTCATTTTGTGTACTTCCACCGCGATATATTTTATCTCATAATATATAAATCT 375
 QY 78 CGAGTTTCATCCATACCATATCTTTTAACAGATTTTATGCTCCGTTTATCGAATA 137
 Db 376 TATATTCGAATTTTTCAGTATTTTATTTTCATATTTTCATCTAATTCATTATTA 435
 QY 138 ACCTAACTCCACTCCGTAGACATGTCATCGTAGGATTCGCTATCAATCGGCAATCG 197
 Db 436 TGTATTTTGTCTTAATCATTTTCAAACTCTCTTTTAAATTTTCAAAAACATCAAAATTTG 495
 QY 198 CAGGACAGTGGCAATACATCTTTACCCCAATGTCGATACGGTTCGCTCGAA 257
 Db 456 AAACATCATCTTCAATATTAATGTCATTTTAAATTCATTATTTCTGTGGCATCTT 555
 QY 258 CTCATTTTC 266
 Db 556 TTCATATTC 564
 RESULT 7

AJ456670/c
 LOCUS 437 bp mRNA linear EST 22-APR-2002
 DEFINITION riken1 Gallus gallus CDNA clone 9d22r1, mRNA sequence.
 ACCESSION AJ456670
 VERSION AJ456670.1 GI:20266766
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AUTHORS Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 TITLE Phasianinae; Gallus.
 JOURNAL 1 (bases 1 to 437)
 COMMENT Gallus gallus bursal lymphocyte EST
 Buerstedde, J.M.
 Unpublished
 Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
 Location/Qualifiers
 1..437
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 /clone="9d22r1"
 /cell_type="bursal lymphocyte"
 /dev_stages="2-3 weeks old"
 /clone_libs="riken1"
 /note="WGS inbred strain"
 BASE COUNT 108 a 99 c 130 g 98 t 2 others
 ORIGIN
 Query Match 14.4%; Score 39; DB 9; Length 437;
 Best Local Similarity 54.2%; Pred. No. 5.9;
 Matches 78; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 62 AAATCATGATTTTCATCGAGTTCAATTCCTATACCATATCTTTAATACGATTTTATGC 121
 Db 336 ATAACAGGAAATCATGACGACCTTGATTCATTCGATTCCTTACTTCTATTTTGCA 277
 QY 122 TCGGTTTATCGAATAACCTACCTTCCTAGCAGACATCGATCGTAGCATTCCT 181
 Db 276 AACCGTAATTCCTCAGTCTATGTCATAGTAGTGGCACTGGCATCAAGGAGTTTCT 217
 QY 182 ATCAACTCGGCAATCGCAGGAACA 205
 Db 216 ATCAGTCTCTTCACCACGCTGACA 193
 RESULT 8
 BQ735322/c
 LOCUS 946 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8510714 NICHD XGC Emb4 Xenopus laevis cDNA clone
 IMAGE:4684253 5', mRNA sequence.
 ACCESSION BQ735322
 VERSION BQ735322.1 GI:21874219
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AUTHORS Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 TITLE Xenopodinae; Xenopus.
 JOURNAL 1 (bases 1 to 946)
 COMMENT NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue procurement: Dr. Igor Dawid
 CDNA library preparation: Life Technologies, Inc.
 CDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Plate: L1AM10568 row: 0 column: 06
 High quality sequence stop: 637.
 Location/Qualifiers

FEATURES
 Source
 1..946
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:4684253"
 /dev_stage="embryo, stage 31-32"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHDD_XGC_Emb4"
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site:1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."
 BASE COUNT 300 a 187 c 199 g 259 t 1 others
 ORIGIN

Query Match 14.1%; Score 38.2; DB 13; Length 946;
 Best Local Similarity 54.7%; Pred. No. 10;
 Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 QY 66 CATTGATTCATCGAGTTCATTCCTATACCATATCTTTTATAACGATTTCCTCCG 125
 Db
 286 CAGAGTTTCGTCACTGTTTGACCTTTCCATTCGCCGAACTTCAATTTGTCAATC 227
 QY 126 GTTATCGAATAACCTAACTCCACTCCGTAGCACATCGTAGCATTCGGTATCA 185
 Db
 226 CAAATTTTCCAACTTAATTTCAATGCTGGTTCAGTACATCAAGGCATTTTCTACTA 167
 QY 186 ACTCGGCAATCGCAGAAC 204
 Db 166 GTTCCTTACCACATAAC 148

RESULT 9
 BM256945
 LOCUS 529 bp mRNA linear EST 17-DEC-2001
 DEFINITION 529410 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BM256945
 VERSION BM256945.1 GI:17892544
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 1 (bases 1 to 527)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett G.L., Heston,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G., Retee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 21180013
 11282978
 CONTACT: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
 PCR Primers

FORWARD: AGGAAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 122 row: K column: 3
 Seq primer: ATTTAGGTGACACTATAG.
 Location/Qualifiers

FEATURES
 Source
 1..527
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 3BOV"
 /note="Vector: pCMV SPORT6; Site:1: NotI; Site 2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."
 BASE COUNT 170 a 74 c 88 g 195 t
 ORIGIN

Query Match 13.9%; Score 37.4; DB 12; Length 527;
 Best Local Similarity 54.8%; Pred. No. 16;
 Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 12 ACGGGAGGCTGCTTTTCTCCCTCTCTGTCGCGCGATTCTCAAATAAAATCATTTGA 71
 Db
 283 ACGGTGAGACTTAGTGGTGTGCTGTATCTCTGATTTTCAAACTGAAAGAGATGT 342
 QY 72 TTTCATCGAAGTTCATTCCTATACCATTCATTTTAAATAACGATTTTCGCGTTTAT 131
 Db
 343 TTTATCTTTAGTTATCTCATTATATAATTAATCAAGTAGATTTTATAAACAGTTTAA 402
 QY 132 CGAATACCTTAACCTT 146
 Db 403 TGCAAAAGTGATTT 417

RESULT 10
 BX443799/c
 LOCUS 1201 bp mRNA linear EST 15-MAY-2003
 DEFINITION BX443799 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CSODM001YA02 5-PRIME, mRNA sequence.
 ACCESSION BX443799
 VERSION BX443799.1 GI:30774179
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1534.f For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODM001BA01Q1&cluster=1534.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/Invitrogen/Corporation/1600ParadiseAvenue/Genoscope/sequenceID%3ACSDM001BA01Q1>.

Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODM001YA02"
 /tissue_type="FETAL LIVER"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL LIVER"

QY 57 AATAAAATCATGATTCATCGAAGTTCATTCCTATACCATTAATCTTTAATAAGATT 116
 Db 606 AATAAAACACAGTGTGTTGGGGTATATTTTTCCTTTTTCITTTTTCACATT 665
 QY 117 TATGCTCGGTTTATCGAATACTTAACCTTCCACTTCGTTAGCAGCATCGTAGGCAT 176
 Db 666 TTTCCTTAGCTTTAGAAATTCCTTGCAGTCTCTTACAAATAACAAGTAAATACCGTTAA 725
 QY 177 TCGCTATCAACTCGCAATCGCAGGAACAGTGTGCGAATACAAATCTTTACACCCAAATGT 236
 Db 726 GTGAAATTACTTGACCAATCATCTTCTTTTTCAGTACAGCTTTTAAAGCAAGTCT 785
 QY 237 TCGATTACGGTTCGCAAACTCAATTCAT 269
 Db 786 TAAAGGCAGTGTAGTCTCAATCATTTTGAAT 818

RESULT 13
 BH934691
 LOCUS
 DEFINITION
 Ode18d03.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
 sequence.
 ACCESSION
 BH934691
 VERSION
 BH934691.1 GI:23414757
 KEYWORDS
 GSS.
 SOURCE
 Brassica oleracea
 ORGANISM
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE
 1 (bases 1 to 688)
 Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash
 W., Rabinowicz,P.D. and Wilson,R.K.
 TITLE
 Whole genome shotgun reads from Brassica oleracea
 JOURNAL
 Unpublished
 COMMENT
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: odel8 row: d column: 03
 Seq primer: -21UpPof forward
 Class: shotgun
 High quality sequence start: 8
 High quality sequence stop: 551.

FEATURES
 source
 Location/Qualifiers
 1..688
 /organism="Brassica oleracea"
 /mol_type="Genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea T0100DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."
 BASE COUNT 246 a 77 c 69 g 296 t
 ORIGIN

Query Match 13.6%; Score 36.8; DB 28; Length 688;
 Best Local Similarity 46.8%; Pred. No. 24;
 Matches 116; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 21 CTTCCTTTCTCCCTCTGTTCGACCGATTCTCAATAAATCAATTCATTCATTCGCA 80
 Db 404 CTTTTATCTTACCACCACTAAGAGACATTTTCAAAATACATTTTTCATTAAGTGA 463
 QY 81 AGTTCATCTATACCAATATCTTTAATAACGATTTTATGCTCGGTTTATCGAATAACC 140
 Db 464 AAAAGACTCTTATACCAATGTATCTCTATATATATAAATTTATTTAATAATATA 523
 QY 141 TAACTTCCACTTCGTTAGCAGTCATCGTAGGCATTCGCTATCACTCGGCAATCGCAG 200

Db 524 TAACACATAATAAAAAATTAATGTTTCGAATATACITTTTCAAAATCGAACTTTTC 583
 QY 201 GAACAGTGCAGTAACAATCTTTTACACCCAAATGTCGATTACGGTTCGCTGAAACTC 260
 Db 584 ATAAATGTTTATTAATGAATTTTTCGAATGTTTATTTTTCATTTTTCAAAATTC 643
 QY 261 AATTTCAA 268
 Db 644 TTTTGGAA 651

RESULT 14
 BE225447/c
 LOCUS
 DEFINITION
 MD0355 Meloidogyne incognita J2 (#MD99-1) Meloidogyne incognita
 cDNA clone 3108 5' similar to unknown Meloidogyne incognita cDNA,
 mRNA sequence.
 BE225447
 BE225447.1 GI:8930683
 EST.
 SOURCE
 Meloidogyne incognita (southern root-knot nematode)
 ORGANISM
 Meloidogyne incognita
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
 REFERENCE
 1 (bases 1 to 801)
 Dautova,M., Gommers,F.J., Bakker,J. and Smart,G.
 TITLE
 5' end expressed sequence tags from Meloidogyne incognita
 J2 cDNA library
 JOURNAL
 Unpublished (2000)
 COMMENT
 Contact: Smart G / Dautova M
 Laboratory of Nematology
 Wageningen University and Research Center
 Box 8123, Binnenhaven 10, 6700 PD Wageningen, The Netherlands
 Tel: 31 317 485 254
 Fax: 31 317 484 254
 Email: Geert.Smart@medew.nema.wau.nl,
 Makedonka.Dautova@medew.nema.wau.nl
 Insert Length: 801 Std Error: 0.00
 Seq primer: T7 promoter primer
 High quality sequence stop: 801.

FEATURES
 source
 Location/Qualifiers
 1..801
 /organism="Meloidogyne incognita"
 /mol_type="mRNA"
 /db_xref="taxon:6306"
 /clone="3108"
 /dev_stage="second stage preparasitic juveniles (J2)"
 /clone_lib="Meloidogyne incognita J2 (#MD99-1)"
 /note="Vector: pMAK1; Site1: Sfi IA; Site2: Sfi IB; cDNA
 was synthesized using SMART III oligo, CDS III oligo d(T
 130 (Clontech) and Superscript II reverse transcriptase
 (Life Technology). cDNA clones were size fractionated and
 directionally ligated in the Sfi IA restriction site at
 5' end and Sfi IB at 3' end of pMAK1. pMAK1 was derived
 from the plasmid pCDNA II (Invitrogen)."
 BASE COUNT 343 a 149 c 208 g 101 t
 ORIGIN

Query Match 13.6%; Score 36.8; DB 10; Length 801;
 Best Local Similarity 54.4%; Pred. No. 24;
 Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 54 TCAATAAAATCATTTGATTTTCATCGAAGTTCATTCCTATACCATTAATCTTTAATAACGA 113
 Db 259 TTAATAAAATTAATGATTTTCCAAAAAATTTTATGGCTTTATATCCCCAAAAA 200
 QY 114 TTTTATGCTCCGGTTTATCGAATACCTAACCTTCACCTTCGTAGCAGCATCGTAGG 173
 Db 199 TCATTACTTATGTTTATTTATTTTATTTTAAATGCAATCTCTTTCATTCGCAATCATC 140
 QY 174 CATTCGCTATCAACTC 189

Sun Nov 16 12:57:03 2003

Db 139 CCTTGTGCTCTCCC 124

RESULT 15
LOCUS CC122282 809 bp DNA linear GSS 16-APR-2003
DEFINITION NDL.20H11.T7 Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION NDL.20H11, genomic survey sequence.
VERSION CC122282
KEYWORDS CC122282.1 GI:29991337
SOURCE GSS.
ORGANISM Aedes aegypti (yellow fever mosquito)
Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
REFERENCE 1 (bases 1 to 809)
AUTHORS Loftus, B., Shetty, J., Knudson, D. and Severson, D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished
COMMENT Other_GSSs: NDL.20H11.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.
FEATURES
source
1..809
Location/Qualifiers
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="NDL.20H11"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pCBAC1; Site 1: Hind III; The library was
prepared from whole body tissue of newly hatched L1 larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"
BASE COUNT 293 a 128 c 68 g 320 t
ORIGIN
Query Match 13.6%; Score 36.8; DB 29; Length 809;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 21 CTGTGTTTCTCCCTTCGTCCGACGATCTCAATAAATCAATGATTCATCGA 80
Db 450 CTCCTTTTCAATTTTCTGTACTTCAATAATACACATAAAATATTGTTTCTCTA 509
QY 81 AGTTCATTCCTATACCATTCCTTAAT 108
Db 510 TTTCGATTTTTCCTCAAAATCTGTAT 537

Search completed: November 15, 2003, 08:02:40
Job time : 1806.19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 42.0547 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-88

Perfect score: 270

Sequence: 1 AATTCTTCGACGGGAGG.....CTCGAAACTCAATTCAATT 270

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA: *

1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.6	13.9	1230025	4	US-09-198-452A-1
C 2	36	13.3	4519	4	US-08-976-259-6
C 3	34.6	12.8	1437	4	US-09-004-838-37
C 4	34.6	12.8	5829	4	US-09-004-838-109
C 5	34.2	12.7	9566	4	US-09-328-352-3038
C 6	34	12.6	7218	1	US-08-232-463-14
C 7	33.2	12.3	14602	1	US-08-597-236-1
C 8	33.2	12.3	14602	1	US-08-746-682A-1
C 9	32.2	11.9	2781	3	US-08-749-522-4
C 10	31.2	11.6	1377	4	US-09-328-352-3423
C 11	31	11.5	2806	4	US-09-463-238-11
C 12	31	11.5	1830121	4	US-09-557-884-1
C 13	31	11.5	1830121	4	US-09-643-990A-1
C 14	30.8	11.4	1152	4	US-09-252-991A-12065
C 15	30.8	11.4	11580	4	US-09-334-220-4
C 16	30.8	11.4	24417	2	US-08-846-762-1
C 17	30.4	11.3	2577	2	US-08-209-521-25
C 18	30.4	11.3	2577	4	US-09-265-503B-135
C 19	30.4	11.3	5361	3	US-08-973-462-2
C 20	30.4	11.3	6152	3	US-08-973-462-1
C 21	30	11.1	501	2	US-08-928-692-56
C 22	30	11.1	501	4	US-09-339-972-56
C 23	30	11.1	1664976	4	US-08-916-421B-1
C 24	29.8	11.0	1465	4	US-09-004-838-100
C 25	29.8	11.0	1465	4	US-09-004-838-33
C 26	29.8	11.0	1474	4	US-09-004-838-36
C 27	29.8	11.0	4211	4	US-09-004-838-106

Sequence 1, Appli
Sequence 39, Appli
Sequence 27, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 21, Appli
Sequence 87, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 8, Appli
Sequence 21, Appli
Sequence 137, App
Sequence 459, App

28 29.8 11.0 1664976 4 US-08-916-421B-1
C 29 29.4 10.9 1380 3 US-08-676-444-39
C 30 29.4 10.9 1439 4 US-09-004-838-27
C 31 29.4 10.9 3760 1 US-08-261-663A-3
C 32 29.4 10.9 3760 5 PCT-US95-07754A-3
C 33 29.4 10.9 3830 1 US-08-261-663A-5
C 34 29.4 10.9 3830 5 PCT-US95-07754A-5
C 35 29.4 10.9 7253 4 US-09-268-347-35
C 36 29.4 10.9 7291 3 US-08-913-942-3
C 37 29.4 10.9 10815 4 US-09-004-838-21
C 38 29.4 10.9 13149 4 US-09-004-838-87
C 39 29.2 10.8 966 1 US-08-257-073-6
C 40 29.2 10.8 1814 1 US-08-097-829-5
C 41 29.2 10.8 1814 1 US-08-577-403-5
C 42 29.2 10.8 3056 4 US-09-708-280-8
C 43 29.2 10.8 3065 2 US-08-209-521-28
C 44 29.2 10.8 3065 4 US-09-265-503B-137
C 45 29 29 592 3 US-09-385-982-459

ALIGNMENTS

RESULT 1
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198/452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(15000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n-a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)


```
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature

Query Match
Best Local Similarity 13.9%; Score 37.6; DB 4; Length 1230025;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 70 GATTTCATCGAAGTTTCATTCTATACCATTTATCTTAATAACGATTTTATGCTCCGGTTT 129
Db 929463 GATGCTTCGGCTCTGAAGCCACACACCATTTATCTGTATATGATGGCCTTGTCCCC 929404

QY 130 ATCGAATTAACCTAATTCATCTCCGTAGACATGATCGTAGGATTCGATTCATCAACTC 189
Db 929403 TCCTAAAGTTTCGATTTCTATTTTCATCGCGCCAGCATCTAAGGAATTTCTCAATCAGTTC 929344

QY 190 GGCAATCGCAGGAACAGTGT 209
Db 929343 TTTAACAACAGAACGGAGT 929324

RESULT 2
US-08-976-259-6
; Sequence 6, Application US/08976259
; Patent No. 6316609
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Choi, Gil H.
; APPLICANT: Welch, Rodney A.
; TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
; Patent No. 6316609
; NUMBER OF SEQUENCES: 142
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,259
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488-0740002/EKS/CBM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: linear
US-08-976-259-6

Query Match
Best Local Similarity 13.38; Score 36; DB 4; Length 4519;
Matches 88; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 47 CCGATTCTCAATAAAAAATCAATTGATTTCATCGAAGTTTCATTCCTATACCATTTATCTTTA 106
```

```
Db 95 CCAATAGTCAACCAACCTCTCTTCAAATTCATCGGTGTGTCATCACTTAATCCATCATCT--- 151
QY 107 ATACGATTTTATGCTCCGGTTTATCGAATAAATCACTTCCACTTCGCTAGCACATGCA 166
Db 152 CTCAGATTAAGAAGATTTCTTTCTTAAATAAATCACTTCGACATATATCAGCATAGGCA 211
QY 167 TCGTAGGATTCGCTATCAACTCGGCAATCGCAGGA 202
Db 212 TCATGAGCATTTTAAATAAATCACTCAAGGCGAGTA 247

RESULT 3
US-09-004-838-37/c
; Sequence 37, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Shen, Kathy
; APPLICANT: Meyers, Blake
; TITLE OF INVENTION: Procedures and Materials for
; TITLE OF INVENTION: Confering Pest Resistance in Plants
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,838
; FILING DATE: 09-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/781,734
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 023070-078810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY:
; LOCATION: 1..1437
; OTHER INFORMATION: /note= "RLG2K"
US-09-004-838-37

Query Match
Best Local Similarity 12.8%; Score 34.6; DB 4; Length 1437;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 44 CGACCGATTCTCAATAAAAAATCAATTGATTTCATCGAAGTTTCATTCCTATACCATTTATCT 103
Db 1422 CCACCAACTCTCATATAAAGCTCTTCAAAGTTTCCAAATTTCTTTAGGACACCATTTATCT 1363
QY 104 TTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTTAATCTTC 148
Db 1362 ATACGAGACCATCAATTTGTCAAATCTAGTACCTTAGCTTC 1318
```

```

RESULT 4
US-09-004-838-109/c
; Sequence 109, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Meyers, Blake
; TITLE OF INVENTION: Procedures and Materials for
; TITLE OF INVENTION: Conferring Pest Resistance in Plants
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/004,838
; FILING DATE: 09-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/781,734
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 023070-078810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5829 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..5829
; OTHER INFORMATION: /note= "RG2K"
US-09-004-838-109

Query Match 12.8%; Score 34.6; DB 4; Length 5829;
Best Local Similarity 58.1%; Pred. No. 0.36;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 44 CGACGATCTCAATAAAATCATTTGATTCATCGAAGTTTCATTCCTATACCATTTCT 103
DB 2678 CCACCAATCTCATATAAGCTTCITCAAGTTTCACCAATTTCTTTAGGACACCATTTCT 2619

QY 104 TTAATAAGATTTTATGCTCCGGTTTATCGAATAACCTAACTTCC 148
DB 2618 ATACGAAGACCATCAACAATTTGTCAAAATCTAGTACCCCTTAGCTCC 2574

RESULT 5
US-09-328-352-3038/c
; Sequence 3038, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3038
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3038

Query Match 12.7%; Score 34.2; DB 4; Length 966;
Best Local Similarity 49.8%; Pred. No. 0.27;
Matches 113; Conservative 0; Mismatches 113; Indels 1; Gaps 1;

QY 28 TTCCTCCCTTCTGTTCCGACCGATCTCAATAAAATCATTTGATTCATCGAAGTTCTAT 87
DB 885 TTCTTCGCTCTTGTGTATAGCAACAATACTTTAGAATCTTTTATACGACCAAGTCTG 826
QY 88 TCCTATACCATTAATCTTTTAATAGGATTTTATGCTCCGGTTTATCGAATAACCTAACTTC 147
DB 825 GATCGCACCAAGAAATACCTACAGCAATGTAAGCTCAGGTGCAACGATTTTAC-CAGTTT 767
QY 148 CACTTCCTAGCAGATGATCGTAGGATTCGCTATCACTCGGCAATCGCAGGAACAGT 207
DB 766 GACCAACTTGAAGTCTGTAGGTACAAAGCCTGCATCAAACTGCTGCAGGTGAACGACTT 707
QY 208 GTGCGAATACAAATCTTTACACCCAAATGTTTCGATTACGGTTGGCTCG 254
DB 706 GTGCTGCACCAAGCTTGTCACTATGGTCAAGTACTTCTGCTGTAG 660

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Hardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match
Best Local Similarity 12.6%; Score 34; DB 1; Length 7218;
Matches 16; Conservative 88; Mismatches 58; Indels 0; Gaps 0;

QY 21 CTGTGTTTCTCCCTCGTTCGACCGATTCTCAATAAAATCATGTTTCATCGA 80
Db 1301 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1360

QY 81 AGTTCATTCTACATATCTTTAATAACAGATTTTATGCTCCGGTTTATCGAATAACC 140
Db 1361 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1420

QY 141 TAACTTCCATTCCTAGACATGATCTAGGATTCGCTA 182
Db 1421 YYYYYYYYYYYYGTACCAAAATCTTCTATCTCTTAACCTA 1462
```

RESULT 7

```

US-08-597-236-1/c
; Sequence 1, Application US/08597236
; Patent No. 5733765
; GENERAL INFORMATION:
; APPLICANT: STINGELE, Francesca
; APPLICANT: MOLLET, Beat
; TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
; TITLE OF INVENTION: EXOPOLYSACCHARIDES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americans
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,236
; FILING DATE:
; CLASSIFICATION: 426
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201669.9
; FILING DATE: 20-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci A., Allan
; REGISTRATION NUMBER: 30256
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1803
; OTHER INFORMATION: /product= "epsa"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1807..2535
; OTHER INFORMATION: /product= "epsB"
; FEATURE:
; NAME/KEY: CDS
```

```

; LOCATION: 2547..3239
; OTHER INFORMATION: /product= "epsC"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3249..3995
; OTHER INFORMATION: /product= "epsD"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4051..4731
; OTHER INFORMATION: /product= "epsE"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4898..5854
; OTHER INFORMATION: /product= "epsF"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6425..7540
; OTHER INFORMATION: /product= "epsG"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7736..8212
; OTHER INFORMATION: /product= "epsH"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8221..9192
; OTHER INFORMATION: /product= "epsI"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9285..10364
; OTHER INFORMATION: /product= "epsJ"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10392..11339
; OTHER INFORMATION: /product= "epsK"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11302..12222
; OTHER INFORMATION: /product= "CDS (eps L) covering CDS
; OTHER INFORMATION: (eps K) on nucleotides 10392-11339"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12233..13651
; OTHER INFORMATION: /product= "epsM"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 13732..14305
; OTHER INFORMATION: /function= "CDS on the
; OTHER INFORMATION: complementary strand"
; OTHER INFORMATION: /product= "Oifz"
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 230..252
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 274..302
; FEATURE:
; NAME/KEY: RBS
; LOCATION: 340..345
; US-08-597-236-1
```

Query Match 12.3%; Score 33.2; DB 1; Length 14602;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```

QY 59 TAAATATCATTCATTCGAGTTCATTCCTATACCATTCCTTATTAACGATTTTA 118
Db 10750 TTAGGATAACTATTTTCAAAAAAATAAATACTATTCCTTATTCGTCATATCA 10691

QY 119 TGTCCGGTTTATCGAATAAACCCTCACTTCCTAGCAGCATCGTAGGCATTC 178
Db 10690 TCTCCAGATAAAAAATGAAATAAGATATTCCTCTGTATCAGTCGACATTC 10631

QY 179 GCTATCACTCGGCAATCCAGGAACAGTGTGCGAATAACAATCTTT 224
```

Db 10630 GCAACATGCTTCTTACCATAGAAAACCTCCCAATGTACATTT 10585

RESULT 8

US-08-746-682A-1/c
 ; Sequence 1, Application US/08746682A
 ; Patent No. 5786184
 ; GENERAL INFORMATION:
 ; APPLICANT: STINGELE, Francesca
 ; APPLICANT: MOLLET, Beat
 ; TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
 ; TITLE OF INVENTION: EXOPOLYSACCHARIDES
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americans
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA: US/08/746,682A
 ; APPLICATION NUMBER: 08/597,236
 ; FILING DATE: 20-JUN-1995
 ; APPLICATION NUMBER: EP 95201669.9
 ; FILING DATE: 20-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci A., Allan
 ; REGISTRATION NUMBER: 30256
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14602 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 352..1803
 ; OTHER INFORMATION: /product= "epsa"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1807..2535
 ; OTHER INFORMATION: /product= "epsb"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2547..3239
 ; OTHER INFORMATION: /product= "epsc"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3249..3995
 ; OTHER INFORMATION: /product= "epsd"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4051..4731
 ; OTHER INFORMATION: /product= "epse"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4898..5854
 ; OTHER INFORMATION: /product= "epsf"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 6425..7540
 ; OTHER INFORMATION: /product= "epsG"

FEATURE:
 NAME/KEY: CDS
 LOCATION: 7736..8212
 OTHER INFORMATION: /product= "epsh"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 8221..9192
 OTHER INFORMATION: /product= "epsi"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 9285..10364
 OTHER INFORMATION: /product= "epsj"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 10392..11339
 OTHER INFORMATION: /product= "epsk"
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 11302..12222
 OTHER INFORMATION: /product= "CDS (eps L) covering CDS
 OTHER INFORMATION: (eps k) on nucleotides 10392-11339"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 12233..13651
 OTHER INFORMATION: /product= "epsm"
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 13732..14305
 OTHER INFORMATION: /function= "CDS on the
 OTHER INFORMATION: complementary strand"
 OTHER INFORMATION: /product= "off2"
 FEATURE:
 NAME/KEY: terminator
 LOCATION: 230..252
 FEATURE:
 NAME/KEY: promoter
 LOCATION: 274..302
 FEATURE:
 NAME/KEY: RBS
 LOCATION: 340..345
 US-08-746-682A-1
 Query Match 12.3%; Score 33.2; DB 1; Length 14602;
 Best Local Similarity 50.0%; Pred No. 1.3;
 Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 59 TAAAAATCATTCATTCGAAAGTTTCCTATACCATTCCTTCCATTCGTCAGACATCGTAGGCATTC 118
 Db 10750 TTAGGATAACTATTTTCAAAAAAATAAATACTATTTTCATTATCTTTGATTGCGCATATCA 10691
 QY 119 TGTCTCGGTTTATCGAATACCTTCCATTCCTTCCATTCGTCAGACATCGTAGGCATTC 178
 Db 10690 TCTCCAGATAAAAAATGAAAAATAGAAATATTCCTCTGTATCAGTCGACATTCCTAATAGC 10631
 QY 179 GCTATCAACTCGCAATCGCAGGAACAGTGTGCGAATACAAATCTTT 224
 Db 10630 GCAACATTCCTTCTTACCATAGAAAAACCTCCCAATGTACATTT 10585
 RESULT 9
 US-08-749-522-4
 ; Sequence 4, Application US/08749522
 ; Patent No. 6096950
 ; GENERAL INFORMATION:
 ; APPLICANT: John, Maliyakal
 ; TITLE OF INVENTION: FIBER-SPECIFIC PROMOTERS
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Quarles & Brady
 ; STREET: 411 East Wisconsin Avenue
 ; CITY: Milwaukee
 ; STATE: WI
 ; COUNTRY: U.S.A.


```
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,522
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 670513.90244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5709
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2781 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-749-522-4

Query Match 11.9%; Score 32.2; DB 3; Length 2781;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 67; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 21 CTGTTTTCCTCCCTTCGTCGACCGATTCTCAATAAAATCATTTGATTCATCGA 80
Db 538 CTGCTTTTCCTTCATTTTTTAAATCACTTCATAATAGATAATAGTAATACCA 597

Qy 81 AGTTCATCTTATACCATATCTTTAATAAGATTTTATGTCGCGTTTATCGAATACC 140
Db 598 GTTCATCGATTTAAATGGATTAGAAATGGTGTGTCCTCTTAAGGACTGAAAAAGA 657

Qy 141 TAACT 145
Db 658 TAACT 662

RESULT 10
US-09-328-352-3423
; Sequence 3423, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3423
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3423

Query Match 11.6%; Score 31.2; DB 4; Length 1377;
Best Local Similarity 52.3%; Pred. No. 2.5;
Matches 69; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 14 GGGGAGGGTGTGTTTTCCTTCCTGTCGACCGATTCTCAATAAAATCATTTGATT 73
Db 1028 GAGATGATCTCATTTTAGTCTGTTTATTTCAGACGAGTCACATAGATACCAATGATA 1087

Qy 74 TCATCGAAGTTCAATCCATACCAATATCTTTAATAACGATTTTATGTCGCGTTTATCG 133
Db 1088 TTGAAGCTGTTCAAAAAGCTTTACAGATGTTTAAAGCCCGATTTGAAGTGTGTTGACGTCG 1147
```

```
Qy 134 AATAACCTAACT 145
Db 1148 TATCACATAACT 1159

RESULT 11
US-09-463-238-11/c
; Sequence 11, Application US/09463238
; Patent No. 6469230
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Alison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (822-826, 2707 and 2797)
; OTHER INFORMATION: n = a or g or c or t
US-09-463-238-11

Query Match 11.5%; Score 31; DB 4; Length 2806;
Best Local Similarity 48.8%; Pred. No. 3.6;
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 1 AATTCTTCGCGACGGGAGGCTGTTTCTCTCCCTCTCTCCGACCGATCTCAATA 60
Db 2796 AATTCGAGGATCGGGTACCATGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2737

Qy 61 AAAATCATTCATTCATCGAAGTTCACTTACCATTATCTTTAATAACGATTTTATG 120
Db 2736 TTTTTCGCGACAGACAGATTCCTACTANTATTTTATATATAAATGCAATTAAT 2677

Qy 121 CTCGGTTTATCGAATAACCTAACTCCACTCCGTCAGACATGCAATC 168
Db 2676 ACAAGGTGATCTTATAAATCATCAGCTTCTCTCCACAAATTTTATC 2629

RESULT 12
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
```

```
/
/
/   SOFTWARE: ASCII Text
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/09/557,884
/   FILING DATE: 25-Apr-2000
/   CLASSIFICATION: <unknown>
/
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: 08/476,102
/   FILING DATE: JUN-5-1995
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Michelle S. Marks
/   REGISTRATION NUMBER: 41,971
/   REFERENCE/DOCKET NUMBER: PB186P3
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: 301-309-8504
/   TELEFAX: 301-309-8439
/   INFORMATION FOR SEQ ID NO: 1:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 1830121 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: double
/   TOPOLOGY: linear
/   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      11.5%; Score 31; DB 4; Length 1830121;
Best Local Similarity 56.3%; Pred. No. 26;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 28 TTCTTCCCTTCGTTCGACGATTCTCAATAAAATCAATTTTCATCGAAGTTTCAT 87
Db 809911 TACATCCCTTAGTACTATGACAGCTAATAATCACCATTAAGATACCAACATTTTTCAT 809852

QY 88 TCCTATACCATATCTTTAATAACGATTTTATGCTCCGGTTTA 130
Db 809851 TGATAAACCATAACTTTAATAAGTAAGTACAGCTGAGCTGTGGTTTA 809809

RESULT 13
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-AUG-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
```

```
/
/
/   NAME: Kenley K. Hoover
/   REGISTRATION NUMBER: 40,302
/   REFERENCE/DOCKET NUMBER: PB186P1C1
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: 301-610-5790
/   TELEFAX: 310-309-8439
/   INFORMATION FOR SEQ ID NO: 1:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 1830121 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: double
/   TOPOLOGY: linear
/   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      11.5%; Score 31; DB 4; Length 1830121;
Best Local Similarity 56.3%; Pred. No. 26;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 28 TTCTTCCCTTCGTTCGACGATTCTCAATAAAATCAATTTTCATCGAAGTTTCAT 87
Db 809911 TACATCCCTTAGTACTATGACAGCTAATAATCACCATTAAGATACCAACATTTTTCAT 809852

QY 88 TCCTATACCATATCTTTAATAACGATTTTATGCTCCGGTTTA 130
Db 809851 TGATAAACCATAACTTTAATAAGTAAGTACAGCTGAGCTGTGGTTTA 809809

RESULT 14
US-09-252-991A-12085/c
; Sequence 12085, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12085
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-12085

Query Match      11.4%; Score 30.8; DB 4; Length 1152;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 77; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 55 CAATATAAAATCAATTCATTCATCGAAGTTTCATTCCTATACCAATATCTTTAATAACGAT 114
Db 423 CAACCTGAAGATCACGAATTTCTCCCAATTTATACATGATGTAATCAAAACCCCA 364

QY 115 TTTATGCTCCGGTTTATCGAATAACCTAACCTCCACTTCCGTAGCACATGATCGTAGGC 174
Db 363 ATATTTTACAATCTTTTCAATATTACTGCTGCAAGTTGCTGTTATCCAGCGCGGTCCAC 304

QY 175 ATTCGCTATCAACTCGGCAATCGGAGGACAGTG 208
Db 303 ATGCAACACCACTGGGCGCAAGCCCAAGATCCTTG 270

RESULT 15
US-09-334-220-4/c
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
```

```
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/OF704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4

Query Match      11.4%; Score 30.8; DB 4; Length 11580;
Best Local Similarity 50.0%; Pred. No. 6.4;
Matches 77; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY      38  CTGTTCCGACCGATTCTCAATAAAATCATTGATTTCATCGAAGTTCATTCCTATACCA 97
Db      3933  CTCATCTGATTCTAGGGTAATCAAAAGCTGCTTCTCATAAAAGTTCGAGGTAAAGTT 3874
QY      98  TTATCTTTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTAACTTCCACTTCCGTA 157
Db      3873  GGATTGATAACTGGGATGATCTGCTTCTGCTTCTCGGACAGAAATGATGATGTCATCGACT 3814
QY      158  GCACATGCATCGTAGGCATTGCTATCAACTCGG 191
Db      3813  GCCCACTGGTCATAGTCTCCCTGAGAACCGG 3780
```

Search completed: November 15, 2003, 08:09:27
Job time : 51.0547 secs

1	270	100.0	270	10	US-09-928-457-88	Sequence 88, Appl
2	121.8	45.1	269	10	US-09-928-457-77	Sequence 77, Appl
3	36	13.3	4519	9	US-09-956-004-6	Sequence 6, Appl
4	34.6	12.8	10151	12	US-10-311-455-2406	Sequence 2406, Ap
5	34.4	12.7	543	9	US-09-864-761-7691	Sequence 7691, Ap
6	34	12.6	15767	12	US-10-311-453-1179	Sequence 1179, Ap
7	34	12.6	15767	12	US-10-240-485-1010	Sequence 105, App
8	33.4	12.4	285	11	US-09-991-936-1627	Sequence 1627, Ap
9	33.4	12.4	285	11	US-09-991-936-1953	Sequence 1953, Ap
10	33.2	12.3	367378	12	US-10-312-841-1	Sequence 1, Appl
11	33	12.2	5360	12	US-10-304-708-66	Sequence 66, Appl
12	33	12.2	5360	12	US-10-311-453-1910	Sequence 1910, Ap
13	32.8	12.1	7511	12	US-10-311-455-1355	Sequence 1355, Ap
14	32.6	12.1	2000	10	US-09-938-842A-4571	Sequence 4571, Ap
15	32.6	12.1	6218	12	US-10-311-453-1339	Sequence 1339, Ap
16	32.6	12.1	6114	12	US-10-032-583-6119	Sequence 6119, Ap


```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 9495..9552
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-2406

Query Match
Best Local Similarity 12.8%; Score 34.6; DB 12; Length 10151;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 41 TTCGACCGATTCTCAATAAAATCATGATTCATCGAAGTTCAATTCCTATACCATTA 100
Db 1152 TCCCAAAACCATCTCTCCTCAACCTTCATAAATAAATACTACTCCTAAACCATCA 1093

QY 101 TCTTTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTAACT 145
Db 1092 TACCTAACAAATTTTATTTATTTTATTTTCTTAATAACAAAT 1048

RESULT 5
US-09-864-761-7691/c
; Sequence 7691, Application US/09864761
; Patent No. US20030048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aconica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,697
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7691
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC016648.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.89
US-09-864-761-7691

Query Match
Best Local Similarity 12.7%; Score 34.4; DB 9; Length 543;
Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 3 TTCTTCGACGGGGAGGCTTGTTTTTCTTCCTCTGTTCCGACCGATTCTCAATAAA 62
Db 508 TTAATCGACAAATGTTTGATAGTTTTCACAATTCCTGTTGTTGATTTCCAGTTTA 449

QY 63 AATCATGATTTCATCGAAGTTTCATTCCTATACCATATCTTTAATAACGATTTTATGCT 122
Db 448 TTCCATGTGGTCAGAGAAGATGCTTGATATTACTTCAAATTTTTCGAATGTTTAAAGAT 389

QY 123 CCGGTTTATCGAATAACCTTA 142
Db 388 TGGTTTTCACATAACATA 369

RESULT 6
US-10-311-455-1179/c
; Sequence 1179, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1179
; LENGTH: 15767
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1179

Query Match
Best Local Similarity 12.6%; Score 34; DB 12; Length 15767;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 41 TTCGACCGATTCTCAATAAAATCATGATTCATCGAAGTTCAATTCCTATACCATTA 100
Db 2231 TTCCAAATTTACTTTTATAACAAAATTTTAACTTAAACCAATAATAATTAATACACTTA 2172

QY 101 TCTTTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTAACTTCCACTTCCGTAGCA 160
Db 2171 ATTITTATTAATATCTACTCTTTCTTTTCCAAATAATCTCACAATCTCTTAAACAACA 2112

```

QY 161 CA 162
Db 2111 AA 2110

RESULT 7

US-10-240-485-105/c
; Sequence 105, Application US/10240485
; Publication No. US20030149327A1
; GENERAL INFORMATION:
; APPLICANT: PISPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCI/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 105
; LENGTH: 15767
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-105

Query Match 12.6%; Score 34; DB 12; Length 15767;

Best Local Similarity 54.9%; Pred. No. 61;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 41 TTCGACCGATTCCTCAATAAATAATCAATGATTCATCGAAGTTTCATCTTACCTATACCATTA 100
Db 2231 TTCCAATTTACTTTTATTAACAAAATTTTAACTTAACCAATAATATTAATTAACACTTA 2172

QY 101 TCTTTAATAACGATTTTATGTCGGTTTATCGAATACCTAACTTCCACTTCCGTAGCA 160
Db 2171 ATTITTATTAATCTACTCTTTCTTTTCCAAATAATCTCACTCTTTTAAACAACA 2112

QY 161 CA 162
Db 2111 AA 2110

RESULT 8

US-09-991-936-1627
; Sequence 1627, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1627
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-1627

Query Match 12.4%; Score 33.4; DB 11; Length 285;
Best Local Similarity 62.7%; Pred. No. 17;
Matches 52; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 61 AAAATCATTGATTTTCATCGAAGTTTCATCTTATACCATTAATCTTTAATAACGATTTATG 120
Db 67 AAAATCATTGATTTATTAATAATCATAGCAAAATTTATCTCGAATATTTATATAGC 126
QY 121 CTCGGTTTATCGAATAACCTAA 143
Db 127 GTCATATTGATCTAATAAAGAA 149

RESULT 9

US-09-991-936-1853
; Sequence 1853, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1853
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-1853

Query Match 12.4%; Score 33.4; DB 11; Length 285;
Best Local Similarity 62.7%; Pred. No. 17;
Matches 52; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 61 AAAATCATTGATTTTCATCGAAGTTTCATCTTATACCATTAATCTTTAATAACGATTTATG 120
Db 67 AAAATCATTGATTTATTAATAATCATAGCAAAATTTATCTCGAATATTTATATAGC 126
QY 121 CTCGGTTTATCGAATAACCTAA 143
Db 127 GTCATATTGATCTAATAAAGAA 149

RESULT 10

US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MEC
; FILE REFERENCE: E01/1206/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778


```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match      12.3%; Score 33.2; DB 12; Length 3673778;
Best Local Similarity 59.6%; Pred. No. 5.8e+02;
Matches 56; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 25 TTTTCTTCCCTCTGTTCCGACCGATTCTCAAAATAAAATCATTCATTCATCGAAGTT 84
Db 2983086 TCTTTTCCCTTTCTTCCACTAACAAATTTCAATAAAACAATTTATTTTACACCCCAATT 2983027

QY 85 CATTCCTATACCATATCTTTTAATACGATTTTA 118
Db 2983026 TCATTAAATATAATAATCTTTTAATAATAATTAA 2982993

RESULT 11
US-10-204-708-66/c
; Sequence 66, Application US/10204708
; Publication No. US20030141852A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 66
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-66

Query Match      12.2%; Score 33; DB 12; Length 5360;
Best Local Similarity 54.5%; Pred. No. 74;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 25 TTTTCTTCCCTCTGTTCCGACCGATTCTCAAAATAAAATCATTCATTCATCGAAGTT 84
Db 3657 TTCTCTTACCTTTTCACTTAACACTATATCCTTAAATAATCCTTACAAATACATA 3598

QY 85 CATTCCTATACCATATCTTTTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTAAC 144
Db 3597 AAAATCTTAACCATTTCTTTATTACAACTACATAATACTCCATTTTATACAAATACCGTAT 3538

QY 145 T 145
Db 3537 T 3537

RESULT 12
US-10-311-455-1255/c
; Sequence 1255, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1255
; LENGTH: 7511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1255

Query Match      12.1%; Score 32.8; DB 12; Length 7511;
Best Local Similarity 46.8%; Pred. No. 96;

```

```

; Sequence 1910, Application US/10311455
; Publication No. US20030143608A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1910
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1910

Query Match      12.2%; Score 33; DB 12; Length 5360;
Best Local Similarity 54.5%; Pred. No. 74;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 25 TTTTCTTCCCTCTGTTCCGACCGATTCTCAAAATAAAATCATTCATTCATCGAAGTT 84
Db 3657 TTCTCTTACCTTTTCACTTAACACTATATCCTTAAATAATCCTTACAAATACATA 3598

QY 85 CATTCCTATACCATATCTTTTAATAACGATTTTATGCTCCGGTTTATCGAATAACCTAAC 144
Db 3597 AAAATCTTAACCATTTCTTTATTACAACTACATAATACTCCATTTTATACAAATACCGTAT 3538

QY 145 T 145
Db 3537 T 3537

RESULT 13
US-10-311-455-1255/c
; Sequence 1255, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1255
; LENGTH: 7511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1255

Query Match      12.1%; Score 32.8; DB 12; Length 7511;
Best Local Similarity 46.8%; Pred. No. 96;

```

Matches 103; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 21 CTGTTTTCCTCCCTTCGTTCGACCGATTCTCAATAAAATCAATGTTTCATCGA 80
Db 1375 CATCATTTTCAACCTTTTACTCTTACATATAATTTAACATAAATATCAATCTACTATAA 1316
QY 81 AGTTCAATCCATACCAATTTCTTTAATAACGATTTTATGCTCGGTTTATCGAATAACC 140
Db 1315 ACCAACTACATATTTTAAAAACAAAATAAACTATAAACTCTTATTTAATACAAA 1256
QY 141 TAACTTCCACTTCCGTAGCAGATGTCGTAGGATTCGTATCAACTCGGCAATCGCAG 200
Db 1255 AAARACCAATACATAACTCAGGCTATAATCCCATCTTTAAACCAACCAATTA 1196
QY 201 GAACAGTGTGCAATACATCTTTACACCCAAATGTTTCA 240
Db 1195 AAATCGCTGAACCAAAATCTCGAATCCAACTTTACAA 1156

RESULT 14

US-09-938-842A-4571
; Sequence 4571, Application US/09938842A
; Parent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4571
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4571

Query Match 12.1%; Score 32.6; DB 10; Length 2000;
Best Local Similarity 54.6%; Pred. No. 63;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 20 GCTGTTTTCCTCCCTTCGTTCGACCGATTCTCAATAAAATCAATGTTTCATCG 79
Db 766 GTTAATATAATTTTCTCCCTCTCAGTGTTCACATAATTTTCAATTTCTCT 825
QY 80 AGTTCAATCCATACCAATTTCTTTAATAACGATTTTATGCTCGGTTTATCGAATAA 138
Db 826 ATTTTCATCTGGGAATAGATCTTTAATCAATGTCATCTCACTGTCACCAAGTAA 884

RESULT 15

US-10-311-455-1239/c
; Sequence 1239, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1239
; LENGTH: 5218
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1239
Query Match 12.1%; Score 32.6; DB 12; Length 5218;
Best Local Similarity 57.3%; Pred. No. 94;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 36 TTCTGTTCCGACCGATTCTCAATAAAATCAATGTTTCATCGAAGTTTCCTATAC 95
Db 3449 TTAAATTAATACTCAATCAAAATAACAAAACAATTAATTTAATTAATCTCATAT 3390
QY 96 CATTATCTTTAATAACGATTTTATGCTCCGGTTTATCGAATAA 138
Db 3389 CATAATATTTAATTTTAAATTAATAAAATTAATCCAAAA 3347

Search completed: November 15, 2003, 08:32:16
Job time : 164.996 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 23:56:24 ; Search time 1181.92 Seconds
(without alignments)
9241.609 Million cell updates/sec
Title: US-09-928-457-89
Perfect score: 267
Sequence: 1 AATTATGAACACACGATCA.....TTACCACTCGGATCAATT 267
Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	267	1	AF169465
2	267	100.0	267	6	AE8918
3	267	100.0	267	6	BD063018
4	241.4	90.4	349061	1	NMA222491
5	238.2	89.2	10522	1	AE002542
6	238.2	89.2	349980	6	AX044034
7	235.8	88.3	339	6	AX04066
8	235.8	88.3	339	6	AX033695
9	36.8	13.8	145242	9	AC008610
10	36.8	13.8	178710	2	AC106766
11	35.6	13.3	178323	2	AC126315
12	35.6	13.3	233045	2	AC132725
13	34.8	13.0	272142	2	AC106944
14	34.6	13.0	827	6	BD147361
15	34.6	13.0	1796	6	AR140098
16	34.6	13.0	1845	6	AR140089
17	34.6	13.0	1845	6	E15865
18	34.6	13.0	1966	6	BD156065
19	34.6	13.0	1966	9	AK001242
20	34.6	13.0	2046	6	AR140090
21	34.6	13.0	2046	6	E15866
22	34.6	13.0	3014	9	AB016789
23	34.6	13.0	3024	6	BD157723
24	34.6	13.0	3024	9	AK022507
25	34.6	13.0	131355	2	AC139655
26	34.6	13.0	254981	10	AL606479
27	34.4	12.9	2000	6	AX655393
28	34.2	12.8	11532	1	AE005744
29	33.8	12.7	125020	9	AF429315
30	33.6	12.6	238046	2	AC095003
31	33.6	12.6	244814	2	AC098223
32	33.4	12.5	174054	2	AC015160
33	33.4	12.5	185497	3	AC010563
34	33.4	12.5	197946	10	AC098722
35	33.4	12.5	201050	1	AE646084
36	33.4	12.5	279088	3	AE003518
37	33.2	12.4	906	1	AB072575
38	33.2	12.4	235918	2	AC127741
39	33.2	12.4	258326	2	AC136583
40	33	12.4	645	9	AB018490
41	33	12.4	3062	9	BC000012
42	32.8	12.3	136212	2	AC139386
43	32.6	12.2	14188	1	AE007237
44	32.6	12.2	172993	9	CNS01RH9
45	32.4	12.1	977	8	AF241266

ALIGNMENTS

RESULT 1
AF169465
LOCUS AF169465
DEFINITION Neisseria meningitidis strain 22491 clone Em029 unknown sequence.
ACCESSION AF169465
VERSION AF169465.1 GI:9754673
KEYWORDS
ORGANISM
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
Ferrin, A., Nassif, X. and Tinsley, C.R.
Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to

```

pathogenic Neisseriae
Unpublished
REFERENCE 2 (bases 1 to 267)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        1..267
        /organism="Neisseria meningitidis"
        /mol_type="genomic DNA"
        /strain="22491"
        /db_xref="taxon:487"
        /clone="Em029"
BASE COUNT 71 a 70 c 61 g 65 t
ORIGIN
Query Match 100.0%; Score 267; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e-67;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
DB 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
QY 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
DB 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
QY 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
DB 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
QY 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
DB 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
QY 241 ATCGCGGTTTACCACATTCGGATGCAATT 267
DB 241 ATCGCGGTTTACCACATTCGGATGCAATT 267

RESULT 3
BD063018 267 bp DNA linear PAT 27-AUG-2002
LOCUS DNA and specific proteins or peptides of the Neisseria meningitidis
DEFINITION species bacteria, method for obtaining them and their biological
APPLICATIONS applications.
ACCESSION BD063018
VERSION BD063018.1 GI:22608621
KEYWORDS JP 2001504684-A/80.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 267)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 80 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/80
PD 10-APR-2001
PF 11-JUL-1997 JP 1998505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..267
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 71 a 70 c 61 g 65 t
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e-67;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
DB 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
QY 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
DB 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
QY 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
DB 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
QY 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
DB 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
QY 241 ATCGCGGTTTACCACATTCGGATGCAATT 267
DB 241 ATCGCGGTTTACCACATTCGGATGCAATT 267

RESULT 2
A68918 267 bp DNA linear PAT 06-MAY-1999
LOCUS Sequence 89 from Patent WO9802547.
DEFINITION SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
ACCESSION A68918
VERSION A68918.1 GI:4759837
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 267)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Merker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
JOURNAL Patent: WO 9802547-A 89 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
COMMENT Other publication FR 2751000 19980116.
FEATURES
    source
        1..267
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 71 a 70 c 61 g 65 t
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e-67;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
DB 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
QY 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
DB 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
QY 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
DB 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
QY 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
DB 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
QY 241 ATCGCGGTTTACCACATTCGGATGCAATT 267
DB 241 ATCGCGGTTTACCACATTCGGATGCAATT 267

```

```

QY 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
DB 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
QY 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
DB 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
QY 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
DB 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
QY 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
DB 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
QY 241 ATCGCGGTTTACCACATTCGGATGCAATT 267
DB 241 ATCGCGGTTTACCACATTCGGATGCAATT 267

RESULT 3
BD063018 267 bp DNA linear PAT 27-AUG-2002
LOCUS DNA and specific proteins or peptides of the Neisseria meningitidis
DEFINITION species bacteria, method for obtaining them and their biological
APPLICATIONS applications.
ACCESSION BD063018
VERSION BD063018.1 GI:22608621
KEYWORDS JP 2001504684-A/80.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 267)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 80 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/80
PD 10-APR-2001
PF 11-JUL-1997 JP 1998505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..267
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 71 a 70 c 61 g 65 t
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e-67;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
DB 1 AATTATGAACACACGCATCATCGTTTCGGCTGCGTTCGTTGCGTTGCGATTAGCAGTTG 60
QY 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
DB 61 CGGCTCAATCAATAATGTAAACGGTTTCGACACAGAACTTCAGNACGTCGCGGTTGC 120
QY 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
DB 121 CTTGGCGTCACCAATGCCGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAA 180
QY 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
DB 181 CTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTATGTATACAGTGTATAGCACA 240
QY 241 ATCGCGGTTTACCACATTCGGATGCAATT 267
DB 241 ATCGCGGTTTACCACATTCGGATGCAATT 267

```

QY 121 CTTGGCGGTACCAATGCGTAAAAATCAGCAACCGCAGCAATGAGGCATACGCATCAA 180
Db 121 CTTGGCGGTACCAATGCGTAAAAATCAGCAACCGCAGCAATGAGGCATACGCATCAA 180
QY 181 CTTTACCGCAATGTGGGTAAACGGGTGACCAATGCTATGTTACCAAGTGTATCAGCACA 240
Db 181 CTTTACCGCAATGTGGGTAAACGGGTGACCAATGCTATGTTACCAAGTGTATCAGCACA 240
QY 241 ATCGGGGTTACCACTTCGGATGCAATT 267
Db 241 ATCGGGGTTACCACTTCGGATGCAATT 267

RESULT 4
LOCUS NMA222491 349061 bp DNA linear BCT 02-SEP-2002
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 2/7.
ACCESSION AL162753 AL157959
VERSION AL162753.2 GI:7379120
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 349061)
Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C.,
Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T.,
Davies, R.M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N.,
Holroyd, S., Jagels, K., Leather, S., Moule, S., Mungall, K.,
Quail, M.A., Rajandream, M.A., Rutherford, K.M., Simmonds, M.,
Skelton, J., Whitehead, S., Spratt, B.G. and Barrall, B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
NATURE 404 (6777), 502-506 (2000)
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
PUBMED 10761919
REFERENCE 2 (bases 1 to 349061)
AUTHORS Parkhill, J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
NOTES
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
LOCATION/Qualifiers
1. 349061
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122597"
/note="serogroup: A"
complement(24..206)
/note="ATR repeat; hmms hit to HMW ATR (1 - 183), score:
310.39"
/label=ATR
209..212
220..681
/gene="NMA0368"
220..681
/gene="NMA0368"
/note="NMA0368"
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=1
/product="putative integral membrane protein"
/protein_id="CAB83669.1"
/db_xref="GI:7379121"
/db_xref="SPTREMBL:Q9JWH7"
/translation="MQEQNRKSPPIVNLVLLVSLVATLSLNVAFVIGNHGSMEGLTV

gene
CDS
LILGSIASLDIRYCAVANYVWLAAILVLLALEKKVVPVHAAPFWGLALVAFSVKAVYV
DEAGNTSDIVRYGAGFYLAFAVAFVAGISGTIFAGKNKKAASAAADGTDKDV"
671..1492
/gene="hemK"
671..1492
/gene="hemK"
/note="NMA0369, hemK, Hemk protein, len: 273 aa; similar
to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
/codon_start=1
/transl_table=1
/product="HemK protein"
/protein_id="CAB83670.1"
/db_xref="GI:7379122"
/db_xref="SPTREMBL:Q9JWH6"
/translation="MTFDKWLGSKLPKNEARMLQYVSEYTRVQLTRGGEBMPDEV
GRWDLAQRLLNGEYVILGAREFYGRFTVPSVLTPRPETELHVAFLARLPEN
RQVADLTGSGAVATVALERPDFAFVRSIDIPPALETARKNAADLCARVEAYGWSWF
DTMPSEKWDIIVSNPPYIENGDKHLDSGLRFEPOIALITDFSDGLSCIRTLAQCAP
DRLAEGGFULLHEHGFDPQGAARVGLAENGSFVETLPDLAQLDRVTLGKYMHLK"
831..840
/gene="hemK"
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1162..1171
/gene="hemK"
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1187..1207
/gene="hemK"
/note="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/note="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1565..1568
1578..2966
/gene="NMA0370"
1578..2966
/note="NMA0370"
/gene="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
P4640 hypothetical protein H10325 (450 aa), fasta scores;
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=1
/product="putative integral membrane protein"
/protein_id="CAB83671.1"
/db_xref="GI:7379123"
/db_xref="SPTREMBL:Q9JWH5"
/translation="MNAVVAIVVIMLVLSRVHVVLSTIGAFVGGAVAGMPLQNI
DAAGVQAGIIPVFNKLEGKALSYALGAFAMATHISGLPQOLAGAVVRKLN
GGMPDSVRSGEGVKKWLLLSILVNGMSONIPIPIAFIMVPLLVNFKIDR
RLIACVITFGLVTTVMFLPYFGALFLNEILLGNTHSAPOLDKVINVMAMAPAL
GMLAGLLAPVHYRPRLYQSNNAADTANADANRPPQSAKSLAAVAIVCFALQ
MTEDSLVGMILGFVFMVLGVNIDKANDVFGEGIKMWVGFIMIAQQGFAAVNA
TGHIOPLVSSMAIFGNSKGMALMLVVLVGLVITMGISGFSFTLPIIAIYVPLCVGL
GSPLIATVAVGTAGALDAGSPASDSTLGTMLGNADQGHDRSDVPTFIHNIP
LLIAGWIAAMVL"
2967..3191
/gene="slyX"
2967..3191
/gene="slyX"
/note="NMA0371, slyX, SLX protein homolog, len: 74 aa;
similar to SLX_HABIN P44759 SLX protein homolog (73 aa),
fasta scores; E(): 0.0017, 33.8% identity in 74 aa
overlap, and SLX_ECOLI P30857 SLX protein (72 aa), fasta
scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=1
/product="SLX protein homolog"
/protein_id="CAB83672.1"

/db_xref="GI:7379124"
/db_xref="SPRMBL:QJWH4"
/translation="MDAVQBFHRIITELEISQALQEDVIAGLNAMVAELRQILDLOQA
QLRLYQMDRNPDAQEPYSLRDEIPPHY"
complement (3271..3423)
/gene="NMA0372"
complement (3271..3423)
/gene="NMA0372"
note="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=11
/product="very hypothetical protein NMA0372"
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="SPRMBL:QJWH3"
/translation="MFRSLPLGSDGFLLAVYEPVGFICLOGKTAVFRSGGRFHRI
KRLUSD"
3562..3571
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement (3586..4356)
/gene="thiP"
complement (3586..4356)
/gene="thiP"
/note="NMA0373, thiP, probable ThiF protein, len: 256 aa;
similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa),
fasta scores; E(): 0.43.1% identity in 246 aa overlap,
and MOEB_ECOLI P12282 molybdoxin biosynthesis MOEB
protein_2 (249 aa), fasta scores; E(): 0.43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdoxin biosynthesis proteins). Contains
pfam match to entry PF00899 ThiF family, ThiF family"
/codon_start=1
/transl_table=11
/product="ThiF protein"
/protein_id="CAB83674.1"
/db_xref="GI:7379126"
/db_xref="SPRMBL:QJWH2"
/translation="MTTEHDNDADFLRYSHILLDEIEGQQLSAHIIWCGG
GTLAAALPYLAASIGITLADSDTVELHRLQVAFDEGDKLKEALADRLHIN
HTVVRTNEKLDGRLTGLVQAADI VLDCDNYATQAVNRACVQAKTPLVSGAAR
REGOLAVRPLDPSPCVACILFGDSASDGICSLFGVPSLVGIIIGTQAAEALKILL
DAGPSGRLLAVRYALGGWQYFLPRNPECPVGAER"
3787..3796
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement (3847..4257)
/gene="thiF"
/note="Pfam match to entry PF00899 ThiF family, ThiF
family, score 186.60, E-value 4.1e-52"
complement (4418..4427)
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4471..7224
/gene="ppc"
4471..7224
/gene="ppc"
/ec_numbers="4.1.1.31"
/note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
917 aa; similar to many e.g. CAPP_RHOFA Q32483
phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
fasta scores; E(): 0.43.3% identity in 928 aa overlap.
Contains 2x Pfam match to entry PF00311 PEPCase,
Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
site motif A (P-loop), and PS00393 Phosphoenolpyruvate
carboxylase active site 2"
/codon_start=1
/transl_table=11
/product="phosphoenolpyruvate carboxylase"
/protein_id="CAB83675.1"

Query Match 90.4%; Score 241.4; DB 1; Length 349061;
Best Local Similarity 98.2%; Pred. No. 1.1e-59;

Matches 266; Conservative 0; Mismatches 1; Indels 4; Gaps 2;
QY 1 AATTATGAACACACGATCATCGTTTCGGCTGCGTTTCGTTGGCATTAGCAGGTTG 60
Db 149147 AATTATGAACACACGATCATCGTTTCGGCTGCGTTTCGTTGGCATTAGCAGGTTG 149206
QY 61 CGGCTCAATCAATATGTAACCGTTTCGACACGAAATTCAGGAACGTGCGCGTTTGC 120
Db 149207 CGGCTCAATCAATATGTAACCGTTTCGACACGAAATTCAGGAACGTGCGCGTTTGC 149266
QY 121 CTTGGCGTCA---CAATCGCGTAATATGTAACCGTTCAGAACCGAGCAATGAAGCATACGCAT 177
Db 149267 CTTGGCGTCAAGCCCAATGCGTAAATATGTAACCGTTCAGAACCGAGCAATGAAGCATACGCAT 149326
QY 178 CAACATTTACCGCAACTGTGGTAAAGCGCGTGA-CAATGCTATGTACCACTGTAATCAG 236
Db 149327 CAACATTTACCGCAACTGTGGTAAAGCGCGTGAAGCAATGCTATGTACCACTGTAATCAG 149386
QY 237 CACAATCGCGTACCACTTCGATGCAATT 267
Db 149387 CACAATCGCGTACCACTTCGATGCAATT 149417

RESULT 5
AE002542/c
LOCUS
DEFINITION
Neisseria meningitidis serogroup B strain MC58 section 184 of 206
of the complete genome.
ACCESSION
AE002542 AE002098
VERSION
AE002542.1 GI:7227188
KEYWORDS
Neisseria meningitidis MC58
SOURCE
Neisseria meningitidis MC58
ORGANISM
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 10522)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
JOURNAL
MEDLINE
20175755
10710307
PUBMED
REFERENCE
2 (bases 1 to 10522)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. .10522
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/db_xref="taxon:122586"
/note="serogroup: B"
complement (72..419)
/gene="NMB1931"
complement (72..419)

FEATURES
source
TITLE
JOURNAL
MEDLINE
20175755
10710307
PUBMED
REFERENCE
2 (bases 1 to 10522)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. .10522
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/db_xref="taxon:122586"
/note="serogroup: B"
complement (72..419)
/gene="NMB1931"
complement (72..419)

gene
CDS

```

/gene="NMB1931"
/notes="hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42260.1"
/db_xref="GI:7227189"
/transl_table="MPVYRTPSRKRLHKMNRQKHKLHGLGEFOTLIFGVGSLHQAE
DAAEAFDRFMDELIAYLAANGFEMAGRLRRFFVPSKIRCPKLRNREKSHYLLAGR
PRRCCHFVRHPAG"
/complement(496..1401)
/gene="NMB1932"
/complement(496..1401)
/gene="NMB1932"
/notes="similar to SP:P00960 GB:J01622 PID:146222
PID:146225 PID:466698 percent identity: 85.67; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Glycyl-tRNA synthetase, alpha chain"
/protein_id="AAF42261.1"
/db_xref="GI:7227190"
/transl_table="MLTFOQIIFKLOTFWADKGTQVLPDMVEVGAGTSHPATCLRAL
GPEPFAAYVOPSRPKDGRYDNPENLQHYQFQVALKPPANIQIYLDLSRELGI
DPKVDIIEFVEDDRENPLGANGLGWELNGMEVTFYFQVGGIDCFPLVGHIT
GIERLAMYQGVENYIDLWAKTLOGNTVTYGDVHQNEVEQSYNFEYSDADWLLRQ
INDYEAQAKRLIAENALALPAYELVLKAGHTNLLDARGAISVTERATYIGIRAL
SRAVQKYVESREKLGFLPMKANA"
/complement(1806..2228)
/gene="NMB1933"
/complement(1806..2228)
/gene="NMB1933"
/notes="similar to GB:L10328 SP:P00832 GB:M25464 GB:V00267
GB:V00311 percent identity: 66.92; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, epsilon subunit"
/protein_id="AAF42262.1"
/db_xref="GI:7227191"
/transl_table="MSINQEVSVSEQKIXSGEATFIVVPTQGELGIYPRHEPIMSL
VRGALATVPEGEKEVLVAVGGVLEVPQDKVTVLADVAVRSMAEMDRARAEEAKAA
EAGISQAKDDKALAAEHKVLAAALQKLTLDYIRSHKK"
/complement(2239..3636)
/gene="NMB1934"
/complement(2239..3636)
/gene="NMB1934"
/notes="similar to SP:P42468 PID:587502 percent identity:
91.52; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, beta subunit"
/protein_id="AAF42263.1"
/db_xref="GI:7227192"
/transl_table="MSQGIQVIIGAVDVVEFPRDMPRVYDALKLDENGLITILEVOQL
LGDGVRAIAMSDDGLKRGMTVNSGTAPITVPVKGTLGRIVDVLGTVPVDEAGPDT
DKSRAIHQAAPKFDLSSTTLELTGIKVIDLLCPRAKGGKVLFGGAGVGKTVNMME
LNNIAKAHAGSLSVFAGVGRTERGNDFYHEMKDSNVLDKVAMYQGMNPPGNRLRV
ALNCTLWAEYPRDEKNGRQDVLFFVDNIYRTLAGTEVSALLGRMPSAVGQPTLL
AENGRQERITSTQTSITSIQAVYVPAADDLTPSPATTFALHLDATVLSRDIASLG
IYPAVDLDDSTSLQVQEHQHYDVARGVQSTLQKYKELRDIILILGMDLSDDEDK
LTVNRARKIQRFLSQPFAHVAEVTGSPGKYVALRDTIAGFKAILNGEYDHLPEQAFYM
VGSIEAEVAKTLN"
/complement(3674..4549)
/gene="NMB1935"
/complement(3674..4549)
/gene="NMB1935"
/notes="similar to GB:L10328 SP:P00837 GB:M25464 GB:V00267
PID:146324 percent identity: 75.43; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11

```

```

/product="ATP synthase F1, gamma subunit"
/protein_id="AAF42264.1"
/db_xref="GI:7227193"
/transl_table="MAVGKEILTIRSVQNTQKITRAMQMVSTSKMRKTQERMLARP
YAEKVRVWHLAQTNTHGIPLESHREIRRVGFILITSDKGLCGGLNANLKFELA
QVOEYRNOQIEEIVFCGSKLMACQSIGLNVASAVNLGDTKEMMLICPLTELFOR
YEKHEIDRIHLVYGVFVNTMRQPRMEVLLPICENVIGDSAPKSPISWEYRPTTALA
ITTELYRRLVSVVYQALSDNMASEQAARWAMKAATDNAGNAIKELRLVYNKSRQAA
LLELYELVAGAAV"
/complement(4574..6121)
/gene="NMB1936"
/complement(4574..6121)
/gene="NMB1936"
/notes="similar to SP:P12985 PID:48337 percent identity:
85.77; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, alpha subunit"
/protein_id="AAF42265.1"
/db_xref="GI:7227194"
/transl_table="MOLNPAEISDLIKAKIENLSVNAEVRTGCTVTSVTDGIVRIHGL
SDAMQCEMLEFPNGTEGLAMNLERDSGAVLGEVEHIKEGDTVCTGTGILSVPGRE
LVGRVVDALGRPIDGKGPINTLTAPIEKIAIGVITARKSVDPQMTGKAISMPVPG
RGRELLIGDROTKTAVALDAIVNKGQIGVICIYVAICQKASIANVRKUEEHGAM
EHTIVVAATASAAALQYIAPYSGCTMGFEFFRDEGDAIVYDDLSKQAVARQSLL
LRPPREAYPGDVFYLRHLLERAAVNEHEVETNGEVKGTGSLTALPIETQA
GDVSAEPTNVI-SITDQIFLETLDFNAGIRPAINAGISVSVGGAAQTKVTKLGGG
IRLALAQYELAAQFQASDLDEATRKQLEHGEVETELMKQKQFSTLTAEMALTWA
INNGSYSDYPAKALAFSEFSLFVRTQHPVELVANASGMSDESEKLEAMKSFK
SSVAYQA"
/complement(6132..6665)
/gene="NMB1937"
/complement(6132..6665)
/gene="NMB1937"
/notes="similar to SP:P12987 PID:48336 percent identity:
59.30; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, delta subunit"
/protein_id="AAF42266.1"
/db_xref="GI:7227195"
/transl_table="NAEPATIAAPYAKALFGLAQEKNOIESWLGLEKLAAYVOEGKV
ASLIRPEPNASEKADILDLVGLDKELKNEVILAGOKRLILPEVYAOQCDITLS
FNHISAVIYSAYPLTDQVGLVQMLNKRFRFSELSKISVEIEPELIGGKIVGVDQVL
DLSVQKUSALYITMTN"
/complement(6670..7140)
/gene="NMB1938"
/complement(6670..7140)
/gene="NMB1938"
/notes="similar to SP:P12989 PID:581814 percent identity:
68.59; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F0, B subunit"
/protein_id="AAF42267.1"
/db_xref="GI:7227196"
/transl_table="WNINATLFAQIIVFGLVWFTMKFWPPIAKALDERAAKVAEGL
AAARXGSDCFEQAEKKVALLAEGNQSEWVANAEKRAKIVVEAKQASSEAARIA
AQAKADVQELPRAREQVAVLAVKGAESILRSEVDASKHAKLDTLQKEL"
/complement(7211..7447)
/gene="NMB1939"
/complement(7211..7447)
/gene="NMB1939"
/notes="similar to SP:P12991 PID:48334 percent identity:
73.75; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F0, C subunit"
/protein_id="AAF42268.1"
/db_xref="GI:7227197"
/transl_table="MGLIAIACGLIVALGALGASIGIAMVGSKYLESSARQPELIGPL
QTKPLFIAGLIDAAFLIGVAIALLFVFNPPAG"
/complement(7504..8370)

```

Gene


```
RESULT 8
AX033695
LOCUS       AX033695
DEFINITION  Sequence 1 from Patent WO0044904.
ACCESSION   AX033695
VERSION     AX033695.1  GI:10280379
KEYWORDS    Neisseria meningitidis
SOURCE      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
ORGANISM    Thonard,J.
REFERENCE   Bas0059 polypeptides from neisseria meningitidis
AUTHORS     PATENT: WO 0044904-A 1 03-AUG-2000;
JOURNAL     SMITHKLINE BEECHAM BIOLOG (BE) ; THONNARD JOELLE (BE)
FEATURES    Location/Qualifiers
            1..339
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /db_xref="taxon:487"
BASE COUNT  93 a 84 c 84 g 78 t
ORIGIN
Query Match      88.3%; Score 235.8; DB 6; Length 339;
Best Local Similarity 97.8%; Pred. No. 4e-59;
Matches 261; Conservative 0; Mismatches 2; Indels 4; Gaps 2;
QY      5 ATGACACACAGCATCATGTTTCGGCTCGCTTGGTTCGTTGGTTCGTTGGCATTAGCAGGTTGGGC 64
Db      1 ATGACACACAGCATCATGTTTCGGCTCGCTTGGTTCGTTGGTTCGTTGGCATTAGCAGGTTGGGC 60
QY      65 TCAATCAATATGTAACCGTTTCGACCCAGAACTTCAGGAACGTCGCGCTTTCGCTTG 124
Db      61 TCAATCAATATGTAACCGTTTCGACCCAGAACTTCAGGAACGTCGCGCTTTCGCTTG 120
QY      125 GCGGTCACCC---AATGCGGTAAATATCAGCAACCGCAGCAATGAGGCATACGATCAAC 181
Db      121 GCGGTCAGCCAAATGCGGTAAATATCAGCAACCGCAGCAATGAGGCATACGATCAAC 180
QY      182 TTACCGCAACTGTGGTGAAGCGGTGA-CCAATGCTATGTTACCACTGTAATCAGCACA 240
Db      181 TTACCGCAACTGTGGTGAAGCGGTGAAGCAATGCTATGTTACCACTGTAATCAGCACA 240
QY      241 ATCGCGGTACCACTTCGATGCAATT 267
Db      241 ATCGCGGTACCACTTCGATGCAATT 267

RESULT 9
AC008610/c
LOCUS       AC008610/c
DEFINITION  Homo sapiens chromosome 5 clone CTB-12904, complete sequence.
ACCESSION   AC008610
VERSION     AC008610.7  GI:15187193
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 145242)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL     Direct Submission
TITLE       Direct Submission
REFERENCE   2 (bases 1 to 145242)
AUTHORS     DOE Joint Genome Institute.
JOURNAL     Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3 (bases 1 to 145242)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission
Submitted (25-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE   4 (bases 1 to 145242)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL     Direct Submission
Submitted (15-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT     On Aug 15, 2001 this sequence version replaced gi:12484308.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.6% of Sequence;
            Estimated Total Number of Errors is 0.5.
            STS Content:
            WI-2865 G04046.
            Location/Qualifiers
            1..145242
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="5"
            /clone="CTB-12904"
BASE COUNT  37848 a 36059 c 36278 g 35057 t
ORIGIN
Query Match      13.8%; Score 36.8; DB 9; Length 145242;
Best Local Similarity 55.5%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY      100 TCAGGAACGTCCGCGCTTTCCTTGGCGGTCCACCAATGCGTAAATAATCAGCAACCGCAG 159
Db      62414 TAAGGACCGCGGCGCTCTCACCGTGGCGTCCACCAACCGTGGGAGCTCTCTCG 62355
QY      160 CAATGAAGGCATACGCATCAACTTTACCGCACTGTGGGTAAAGCGGTGACCAATGCTAT 219
Db      62354 CGAGACCGACTGCGGCGTCCACATCAACGACGCGCGAGATCGGCGTGGCCAGCACCA 62295
QY      220 GTTACCAG 227
Db      62294 GTTAGGAG 62287

AC106766     178710 bp  DNA  linear  HTG 12-JAN-2002
LOCUS       AC106766/c
DEFINITION  Homo sapiens chromosome 5 clone RP11-252M5, WORKING DRAFT SEQUENCE,
            11 unordered pieces.
ACCESSION   AC106766
VERSION     AC106766.1  GI:18139316
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 178710)
AUTHORS     DOE Joint Genome Institute.
JOURNAL     Direct Submission
TITLE       Direct Submission
REFERENCE   2 (bases 1 to 178710)
AUTHORS     DOE Joint Genome Institute.
JOURNAL     Submitted (12-JAN-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Center Project Name: 503426
            Center clone name: RPCI-11_252M5
```

```

-----
Summary Statistics
Consensus quality: 168617 bases at least Q40
Consensus quality: 170711 bases at least Q30
Consensus quality: 171417 bases at least Q20
Estimated insert size: 178230; agarose-fp estimation
Estimated insert size: 177710; sum-of-contigs estimation
Quality coverage: 10.04 in Q20 bases; agarose-fp estimation
Quality coverage: 10.07 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1712: contig of 1712 bp in length
* 1713: gap of unknown length
* 1812: gap of unknown length
* 1813: contig of 1822 bp in length
* 3504: gap of unknown length
* 3505: gap of unknown length
* 4916: contig of 1312 bp in length
* 5016: gap of unknown length
* 5017: contig of 2581 bp in length
* 7597: gap of unknown length
* 7598: gap of unknown length
* 7698: contig of 5915 bp in length
* 13612: gap of unknown length
* 13712: gap of unknown length
* 31775: contig of 18063 bp in length
* 31875: gap of unknown length
* 31876: gap of unknown length
* 53671: contig of 21796 bp in length
* 53672: gap of unknown length
* 53771: gap of unknown length
* 53772: contig of 20004 bp in length
* 73776: gap of unknown length
* 73775: gap of unknown length
* 95412: contig of 21537 bp in length
* 95412: gap of unknown length
* 95512: gap of unknown length
* 95513: contig of 30824 bp in length
* 126336: gap of unknown length
* 126337: gap of unknown length
* 126437: 178710: contig of 52274 bp in length.
*
FEATURES
Location/Qualifiers
1. 178710
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-252M5"
/clone_lib="RPCI human BAC library 11"
BASE COUNT 45146 a 44147 c 43668 g 44699 t 1050 others
ORIGIN
Query Match 13.8%; Score 36.8; DB 2; Length 178710;
Best Local Similarity 55.5%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 100 TCAGAACGTCGCGGTTGCTTGGCGTCACCAATGCCGTAAATATCAGCAACCCGAG 159
DB 127083 TAAGACCGCGCGGCTCTCACCGTGGCGTCACCAACACCGCGGCGAGTCCATCTCTCG 127024
QY 160 CAATGAAGGCATACGCATCAACTTACCGCAACTGTGGTAAGCGCGTACCAATGCTAT 219
DB 127023 CGAGACCGACTGCGCGTCCACATCAACGCGGCGGAGATCGCGTGGCCAGCAACAA 126964
QY 220 GTTACCAG 227
DB 126963 GTTAGGAG 126956
RESULT 11
AC126315 178233 bp DNA linear HTG 19-NOV-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-409C19, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
ACCESSION
AC126315
VERSION
AC126315.4 GI:25080517

```

KEYWORDS SOURCE ORGANISM

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 178323)
Muzny, D., Marie, M., Mettler, M., Lee, A., Adams, C., Alder, J.,
Allen, C., Allen, H., Altschuld, S., Amin, A., Anguiano, D.,
Anyalechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Balek, D., Bandaru, A., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idelbird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovacs, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Louie, H., Lozano, R., Lu, X., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwankwelu, O., Okwionu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Reiter, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S., Shen, H.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shetty, J.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL
Unpublished
2 (bases 1 to 178323)
Worley, K.C.
Direct Submission
Submitted (05-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178323)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23195448.
The sequence in this assembly is a combination of BAC based reads
and whole genome sequencing reads assembled using Atlas

REFERENCE

AC126315 178233 bp DNA linear HTG 19-NOV-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-409C19, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
ACCESSION
AC126315
VERSION
AC126315.4 GI:25080517

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZ08
Center clone name: CH230-409C19
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 145050 bases at least Q40
Consensus quality: 147946 bases at least Q30
Consensus quality: 149916 bases at least Q20
Estimated insert size: 150833; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 171060: contig of 171060 bp in length
* 171061 171160: gap of unknown length
* 171161 173957: contig of 2797 bp in length
* 173958 174057: gap of unknown length
* 174058 178323: contig of 4266 bp in length.

Location/Qualifiers
1..178323
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-409C19"
1..1376
/note="wgs end extension"
clone_end:T7"
3965..4975
/note="clone boundary"
clone_end:T7
site:
end sequence:BZ249410"
complement(167238..168180)
/note="clone boundary"
clone_end:Sp6
site:
end sequence:RXB8A22TV"
41076 a 34387 c 34604 g 41395 t 26861 others

BASE COUNT 41076 a 34387 c 34604 g 41395 t 26861 others
ORIGIN
Query Match 13.3%; Score 35.6; DB 2; Length 178323;
- Best Local Similarity 50.6%; Pred. No. 24;
Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 50 TTACGAGTTGGGCTCAATCAATGTAACCGTTTCGACCGAGAACTTCAGGAACGT 109
Db 59751 TTGAATCGTAAAGCAAGATCTTGCTATATCTGTGTTAAACCGAGAACTTTAGAAACTT 59810

QY 110 GCCGGTTTGGTGGGCGTACCAGTCCCGTAAATACCAACCCGACCAATGAAGGC 169
Unpublished
2 (bases 1 to 233045)

Db 59811 TAGAAACTTCCCTGGAGCCTTAAGATTCTTCTGTATTATCACCAGCAGACTAGATACGCA 59870
QY 170 ATACGGATCAACTTACCCCACTGGGTAAAGCGGTGACCAATGCTAT 219
Db 59871 TGTATAGAACCTTTAGACCAATTGGAACACACCTTAAACAAATGCTAT 59920
RESULT 12
AC132725 233045 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-351P24, WORKING DRAFT SEQUENCE, 5
unordered pieces.
AC132725
AC132725.3 GI:25089932
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 233045)
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, W., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davilla, W. L., Davis, C., Davy-Carroll, L., De Andrada, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, X., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, N., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwackeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quitoz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 233045)

TITLE
JOURNAL
REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (03-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 233045)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23811366.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUQI
 Center clone name: CH230-351P24
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 205239 bases at least Q40
 Consensus quality: 207848 bases at least Q30
 Consensus quality: 209231 bases at least Q20
 Estimated insert size: 208878; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: this is a 'working draft' sequence. It currently
 consists of 6 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 61478: contig of 61478 bp in length
 61479 61578: gap of unknown length
 61579 68503: contig of 6925 bp in length
 68504 68603: gap of unknown length
 68604 154066: contig of 85463 bp in length
 154067 154166: gap of unknown length
 154167 230594: contig of 76428 bp in length
 230595 231905: gap of unknown length
 231906 232005: contig of 1211 bp in length
 232006 233045: contig of 1040 bp in length.

 Location/Qualifiers
 1.233045
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-351P24"
 2676. 3550
 misc_feature
 /note="clone boundary
 clone_end:Sp6
 site:
 end sequence:B225033"
 53991. 54868
 misc_feature

/note="clone boundary
 clone_end:T7
 end sequence:B225031"
 57920. 60327
 misc_feature
 /note="wgs_end_extension.
 clone_end:T7"
 60378. 61478
 misc_feature
 /note="wgs_end_extension
 clone_end:T7"
 154167. 155279
 misc_feature
 /note="wgs_end_extension
 clone_end:T7"
 58983 a 47651 c 47755 g 55805 t 22851 others
 BASE COUNT
 ORIGIN

Query Match 13.3%; Score 35.6; DB 2; Length 233045;
 Best Local Similarity 50.16%; Pred. No. 24;
 Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 50 TTAGCAGGTTCGGCTCAATCAATGTAACCGTTCCGACCCAGAACTTCAGGAAGT 109
 |||||
 Db 10340 TTGAATGTTAGCAAGAATCTTCTATATCTGTTTAAACCGAACTTTAGAACTT 10399
 |||||

QY 110 GCGCGTTTTCCTGGCGTCCCAATGCCGTAATAAATCAGCAACCGCAGCAATGAAGGC 169
 |||||
 Db 10400 TAGAAACTTCCTGGAGCCTTAAGATTCTTCTGTATACCGACACATAGATACGCA 10459
 |||||

QY 170 ATACGCATCACTTTACCGCACTGCGGTAGCGCGTGACCAATGCTAT 219
 |||||

Db 10460 TGTCTAGAACCTTTAGACCAATTTGGAACACACCTTAAACAAATGCTAT 10509
 |||||

RESULT 13
 AC106944/c
 LOCUS
 DEFINITION
 AC106944
 AC106944.4 GI:22857067
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 272142)
 Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S,, Adams,C,, Alder,J.,,
 Allen,C,, Allen,H,, Albrooks,S,, Amin,A., Anguiano,D.,,
 Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,,
 Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,,
 Harvey,I., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,,
 Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,M.,,
 Hollins,B., Howells,S., Huiyk,S., Hume,J., Idlebird,D., Jackson,A.,,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,,
 Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaekelemeh, O., Okwuon, G., Olarnpungagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quatro, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Staimis, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D. von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 272142)
 Worley, K.C.
 Direct Submission
 Submitted (14-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 272142)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 14, 2002 this sequence version replaced gi:21737096.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GONT
 Center clone name: CH230-3718
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 220398 bases at least Q40
 Consensus quality: 224532 bases at least Q30
 Consensus quality: 226624 bases at least Q20
 Estimated insert size: 243846; sum-of-contigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 5 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 264441: contig of 264441 bp in length
 264442: gap of unknown length
 264541: contig of 1119 bp in length
 265660: gap of unknown length
 265661: gap of unknown length
 265662: gap of unknown length
 267487: gap of unknown length
 267587: gap of unknown length
 268752: gap of unknown length
 268753: gap of unknown length
 268853: contig of 3290 bp in length.

FEATURES
 Location/Qualifiers
 1..272142
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-3718"
 1..1619
 /note="wgs contig"
 260733..261787
 /note="wgs contig"
 misc_feature
 misc_feature
 BASE COUNT 71571 a 43698 c 43504 g 69419 t 43950 others
 ORIGIN

Query Match 13.0%; Score 34.8; DB 2; Length 272142;
 Best Local Similarity 48.9%; Pred. No. 42;
 Matches 93; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 49 ATTACGAGTTGGCGTCAATCAATTAATGTAACCGTTCCGACCAAGAACTTCAGGAAG 108
 Db 106761 ATTACCATGTTGAGTTTCATCTTATGTGATGACTTAATTCATAATTTACTGAAA 106702
 QY 109 TGGCGGGTTGCTTGGCGTCACCAATCCGTAATAATCAGCAACCGCAGCAATCAAGG 168
 Db 106701 TTCTATTGACTCTAGTTTACACAGATGTGATTAAGAAATACCATATCTTTT 106642
 QY 169 CATACCATCAACTTTTACCGCAACTGTGGTAAGCGCGTGACCAATGTATGTACCACT 228
 Db 106641 TATTACATTCATTCTTCTCAAAAATTTAGGACGAGCTTGAGCAGTACTTGTAACAAGA 106582
 QY 229 GTATCAGCA 238
 Db 106581 GTGATAAGCA 106572

RESULT 14
 LOCUS BD147361 827 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD147361
 VERSION BD147361.1 GI:27853119
 KEYWORDS JP 2002191363-A/2204.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 827)
 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 2204 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/2204
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUO OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 : Search time 172.322 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-89
Perfect score: 267
Sequence: 1 AATTATGACACACGATCA.....TTACCACTTCGATGCAATT 267

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

N. gonorrhoeae nuc
N. gonorrhoeae nuc
Human cDNA clone (H
Human cDNA #506 di
Glutamine:fructose
Glutamine:fructose
Human cDNA sequenc
Glutamine:fructose
Human GPAT II codi
GPAT related nucle
Nucleotide sequenc
Human cDNA sequenc
GPAT related nucle
Human cDNA #895 di
DNA encoding novel
GPAT related nucle
Listeria monocytog
Listeria monocytog
Oligonucleotide fo
Oligonucleotide fo
Bovine embryonic g
Genomic sequence o
Human ovarian can
Listeria monocytog
Drosophila melanog
Drosophila melanog
Listeria monocytog
Human polynucleoti
Human polynucleoti
Listeria monocytog
Transporters and i
Transporters and i
Human membrane tra
Mouse ischaemic co
DNA encoding novel
DNA encoding novel
Mouse GPAT II codi

ALIGNMENTS

RESULT 1									
AAV03597									
ID	AAV03597	standard; DNA; 267 BP.							
XX	AAV03597;								
XX									
XX									
XX	22-OCT-1998	(first entry)							
XX									
XX	Neisseria meningitidis	DNA sequence E29.							
XX									
XX	N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;								
KW	region 3; pathogenicity; blood-brain barrier; diagnosis; infection;								
KW	meningitis; ss.								
XX									
OS	Neisseria meningitidis.								
XX									
XX	WO9802547-A2.								
PN									
XX									
PD	22-JAN-1998.								
XX									
PF	11-JUL-1997;	97WO-FR01295.							
XX									
PR	12-JUL-1996;	96FR-0008768.							
XX									
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.								
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.								
PA	(SMIK) SMITHKLINE BEECHAM.								
XX									
PI	Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;								
PI	Vinals C;								
XX									

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	267	19 AAV03597	Neisseria meningit
2	267	100.0	267	19 AAA15394	Genomic fragment o
C 3	238.2	89.2	46593	21 AAA81456	N. meningitidis pa
C 4	238.2	89.2	34980	21 AAF21612	Neisseria meningit
C 5	238.2	89.2	837096	21 AAA81489	N. meningitidis pa
6	235.8	88.3	339	21 AAA76029	Neisseria meningit
7	235.8	88.3	339	21 AAA15300	DNA encoding a pol
8	226.2	84.7	12077	21 AAA81734	N. meningitidis pa

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

DR WPI; 1998-110594/10.
 XX Genes present in *Neisseria meningitidis* but not other *Neisseria*
 PT species - and related host cells, RNA, anti-sense sequences,
 PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*
 PT meningitidis infection and in protective vaccines
 XX
 PS Example 4; Page 125; 150pp; French.

XX AA03575-606 represent sequences that are present in *Neisseria*
 CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the
 CC genes involved in biosynthesis of the capsule polysaccharide, *frpA* or *C*,
 CC *opc*, *porA*, *rotamase*, sequence *IC1106*, *iga* protease, *pillin*, *pilC*,
 CC proteins which bind transferrin and opacity proteins. The DNA sequences
 CC are responsible for the differences in pathogenicity between *N.*
 CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that
 CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences
 CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*
 CC *lactamica*, are responsible for colonisation and penetration of the
 CC mucosa. The DNA sequences can be used to produce probes and primers, and
 CC antibodies produced against the encoded proteins are used in standard
 CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*
 CC infection, particularly meningitis.

XX Sequence 267 BP; 71 A; 70 C; 61 G; 65 T; 0 other;

Query Match 100.0%; Score 267; DB 19; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.5e-80;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATTATGACACAGCATCATCGTTTCGGTGGCGTTTCGTTGGCATTAGCAGTTG 60
 DB 1 AATTATGACACAGCATCATCGTTTCGGTGGCGTTTCGTTGGCATTAGCAGTTG 60
 QY 61 CGGCTCAATCAATTAATGTAACCGTTTCGACACAGAACTTCAGGAACGTGCCGCTTGC 120
 DB 61 CGGCTCAATCAATTAATGTAACCGTTTCGACACAGAACTTCAGGAACGTGCCGCTTGC 120
 QY 121 CTGGGCGTCAACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAA 180
 DB 121 CTGGGCGTCAACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAA 180
 QY 181 CTTTACCCCAACTGTGGTGAAGCGGTGACCAATGCTATGTTACCACTGTAAATCAGCACA 240
 DB 181 CTTTACCCCAACTGTGGTGAAGCGGTGACCAATGCTATGTTACCACTGTAAATCAGCACA 240
 QY 241 ATCGCGTTTACCACTTCGATGCAATT 267
 DB 241 ATCGCGTTTACCACTTCGATGCAATT 267

RESULT 2

AAA15394
 ID AAA15394 standard; DNA; 267 BP.
 XX
 AC AAA15394;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Genomic fragment of *Neisseria meningitidis* Z2491.
 XX
 KW Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO200026375-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 28-OCT-1999; 99WO-FR02643.
 XX
 PR 30-OCT-1998; 98FR-0013693.
 XX

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 XX
 DR WPI; 2000-365622/31.

XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 PT or preventative vaccines and for diagnosis
 PT
 PS Claim 1; Page 43; 187pp; French.

XX The present sequence represents a genomic fragment of *Neisseria*
 CC meningitidis. The specification describes proteins that are specific
 CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,
 CC or their antigenic fragments, are used in vaccines to treat or protect
 CC against *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequences are also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.
 CC note: the present sequence is included in claim 1, but it is those
 CC sequences that do not include the present sequence that are actually
 CC claimed.

XX Sequence 267 BP; 71 A; 70 C; 61 G; 65 T; 0 other;

Query Match 100.0%; Score 267; DB 21; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.5e-80;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATTATGACACACGCATCATCGTTTCGGTGGCGTTTCGTTGGCATTAGCAGTTG 60
 DB 1 AATTATGACACACGCATCATCGTTTCGGTGGCGTTTCGTTGGCATTAGCAGTTG 60
 QY 61 CGGCTCAATCAATTAATGTAACCGTTTCGACACAGAACTTCAGGAACGTGCCGCTTGC 120
 DB 61 CGGCTCAATCAATTAATGTAACCGTTTCGACACAGAACTTCAGGAACGTGCCGCTTGC 120
 QY 121 CTGGGCGTCAACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAA 180
 DB 121 CTGGGCGTCAACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAA 180
 QY 181 CTTTACCCCAACTGTGGTGAAGCGGTGACCAATGCTATGTTACCACTGTAAATCAGCACA 240
 DB 181 CTTTACCCCAACTGTGGTGAAGCGGTGACCAATGCTATGTTACCACTGTAAATCAGCACA 240
 QY 241 ATCGCGTTTACCACTTCGATGCAATT 267
 DB 241 ATCGCGTTTACCACTTCGATGCAATT 267

RESULT 3

AAA81456/C
 ID AAA81456 standard; DNA; 46593 BP.
 XX
 AC AAA81456;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE *N. meningitidis* partial DNA sequence gnm_4 SEQ ID NO:4.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW *Meningococcus B*; MenB; ds.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO200022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23573.
 XX


```

XX 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX (CHIR ) CHIRON CORP.
PA
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
PI WPI; 2000-318079/27.
DR
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea
XX
XX Claim 7; Page 260-274; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MerB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
XX Sequence 46593 BP; 11355 A; 13195 C; 11355 G; 10687 T; 1 other;
SQ
Query Match 89.2%; Score 238.2; DB 21; Length 46593;
Best Local Similarity 97.4%; Pred. No. 1.3e-69;
Matches 264; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 1 AATTATGACACACGCATCATCGTTTCGGTTCGCTGCGTTTCGTTGCGATTAGCAGTTG 60
Db 46140 AATTATGACACACGCATCATCGTTTCGGTTCGCTGCGTTTCGTTGCGATTAGCAGTTG 46081
QY 61 CGGCTCAATCAATATGTAACCGTTTCGACCCAGAAACTTCAGGAACGTCCGCGTTTGC 120
Db 46080 CGGCTCAATCAATATGTAACCGTTTCGACTAGAACTTCAGGAACGTCCGCGTTTGC 46021
QY 121 CTTGGCGGTCA---CCAATGCGGTAAAAATCAGCAACCGCAGCAATGAAGCATACGCAT 177
Db 46020 CTTGGCGGTGAGCCCAATGCGGTAAAAATCAGCAACCGCAGCAATGAAGCATACGCAT 45961
QY 178 CAACCTTTACCGCACTGTGGGTAGCGGTGA-CCAATGCTATGTTACCACTGAATCAG 236
Db 45960 CAACCTTTACCGCACTGTGGGTAGCGGTGAGCAATGCTATGTTACCACTGAATCAG 45901
QY 237 CACAATCGCGGTACCACTCCGATGCAATT 267
Db 45900 CACAATCGCGGTACCACTCCGATGCAATT 45870

RESULT 4
AAF21612/c

```

```

ID AAF21612 standard; DNA; 349980 BP.
XX
AC AAF21612;
XX
DT 13-MAR-2001 (first entry)
XX
XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
DE
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
XX Neisseria meningitidis.
OS
XX WO2000066791-A1.
PN
XX 09-NOV-2000.
PD
XX
XX 08-MAR-2000; 2000WO-US05928.
PF
XX 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
XX WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -
XX Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
XX Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;
SQ
Query Match 89.2%; Score 238.2; DB 21; Length 349980;
Best Local Similarity 97.4%; Pred. No. 2.9e-69;
Matches 264; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 1 AATTATGACACACGCATCATCGTTTCGGTTCGCTGCGTTTCGTTGCGATTAGCAGTTG 60
Db 239260 AATTATGACACACGCATCATCGTTTCGGTTCGCTGCGTTTCGATTAGCAGTTG 239201
QY 61 CGGCTCAATCAATATGTAACCGTTTCGACCCAGAAACTTCAGGAACGTCCGCGTTTGC 120

```

Db 239200 CGGCTCAATCAATGTAACCGTTTCGACTAGAACTTCAGGAACGTGCGCGTTTCG 239141

QY 121 TTGGGGCTCA---CCAATGCGGTAAATAATCAGCAACGCGAGCAATGAAGCATACGCAT 177

Db 239140 CTTGGGCTCAGCCCAATGCGGTAAATAATCAGCAACGCGAGCAATGAAGCATACGCAT 239081

QY 178 CAACCTTTACCGCAACTGCGGTAAAGCGGTGA---CCAATGCTATGTTACCACTGTAATCAG 236

Db 239080 CAACCTTTACCGCAACTGCGGTAAAGCGGTGA---CCAATGCTATGTTACCACTGTAATCAG 239021

QY 237 CACAATGCGGTTTACCACTTCGATGCAATT 267

Db 239020 CACAATGCGGTTTACCACTTCGATGCAATT 238990

RESULT 5

AAA81489/c

ID AAA81489 standard; DNA; 837096 BP.

XX

AC AAA81489;

XX

DT 04-DEC-2000 (first entry)

XX

DE N. meningitidis partial DNA sequence gmm_37 SEQ ID NO:37.

XX

KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

KW Meningococcus B; MenB; ds.

XX

OS Neisseria meningitidis.

XX

PN WO200022430-A2.

PD

PF 08-OCT-1999; 99WO-US23573.

PR 09-OCT-1998; 98US-0103794.

PR 30-APR-1999; 99US-0132068.

XX

PA (CHIR) CHIRON CORP.

XX

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V;

PI Rappignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

PI Rappignani R, Pizza M;

XX

DR WPI; 2000-318079/27.

XX

PT Isolated nucleotide sequences of Neisseria meningitidis which can be

PT used in the diagnosis and treatment of N. meningitidis infection and

PT other Neisserial infections, for example, N.gonorrhoea -

XX

PS Claim 7; Page 629-865; 1760pp; English.

XX

CC The present invention describes methods of obtaining immunogenic

CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414

CC represent specifically claimed Neisseria meningitidis genomic DNA

CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to

CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to

CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF

CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The

CC composition can be used as a medicament (or in the manufacture of a

CC medicament) for treating, preventing or diagnosing infection due to

CC Neisserial bacteria. For example, some of the identified proteins could

CC be components of vaccines against Meningococcus B; against all serotypes;

CC and/or against all pathogenic Neisseriae. Identification of sequences

CC from the bacterium will also facilitate production of biological probes,

CC particularly organism-specific probes. Attempts to make efficacious

CC Meningococcus B vaccines have failed mainly due to antigen tolerance.

CC Multivalent vaccines have also been tried but none have successfully

CC overcome antigenic variability. The provision of further, complete

CC sequences may provide an opportunity to identify secreted or surface

CC exposed proteins that may be presumed targets for the immune system and

CC which are not antigenically variable or at least more conserved than

CC other more variable regions.

XX

SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Query Match 89.2%; Score 238.2; DB 21; Length 837096;

Best Local Similarity 97.4%; Pred No. 4.2e-69;

Matches 264; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 1 AATTATGAACACACGCATCATCGTTTCGCTCGCTTCGTTGGCATTAGCAGGTTG 60

Db 604031 AATTATGAACACACGCATCATCGTTTCGCTCGCTTCGTTGGCATTAGCAGGTTG 603972

QY 61 CGGCTCAATCAATGTAACCGTTTCCGACACAGAACTTCAGGAACGTGCGCGTTTGC 120

Db 603971 CGGCTCAATCAATGTAACCGTTTCCGACACAGAACTTCAGGAACGTGCGCGTTTGC 603912

QY 121 CTTGGGCGTCA---CCAATGCGGTAAATAATCAGCAACGCGCAATGAAGCATACGCAT 177

Db 603911 CTTGGGCGTCA---CCAATGCGGTAAATAATCAGCAACGCGCAATGAAGCATACGCAT 603852

QY 178 CAACCTTTACCGCAACTGCGGTAAAGCGGTGA---CCAATGCTATGTTACCACTGTAATCAG 236

Db 603851 CAACCTTTACCGCAACTGCGGTAAAGCGGTGA---CCAATGCTATGTTACCACTGTAATCAG 603792

QY 237 CACAATGCGGCTTACCACTTCGATGCAATT 267

Db 603791 CACAATGCGGCTTACCACTTCGATGCAATT 603761

RESULT 6

AAA76029

ID AAA76029 standard; DNA; 339 BP.

XX

AC AAA76029;

XX

DT 19-JAN-2001 (first entry)

XX

DE Neisseria meningitidis BASB059 gene.

XX

KW BASB059; respiratory tract infection; invasive bacterial disease;

KW bacteraemia; meningitis; ds.

XX

OS Neisseria meningitidis.

XX

FN Key Location/Qualifiers

FT 1.339

FT CDS /*tag= a

FT /product= "BASB059"

XX

PN WO200044904-A1.

XX

PD 03-AUG-2000.

XX

PF 25-JAN-2000; 2000WO-EP00561.

XX

PR 29-JAN-1999; 99GB-0002070.

XX

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX

PI Thonnard J;

XX

DR WPI; 2000-505978/45.

DR P-PSDB; AAB21107.

XX

PT New isolated polypeptide from Neisseria meningitidis is useful for

PT detection and treatment of N. meningitidis infection -

XX

PS Claim 7; Page 61; 77pp; English.

XX PS Claim 7; Page 1565-1568; 1760pp; English.

XX CC The present invention describes methods of obtaining immunogenic

XX CC proteins from *Neisseria* genomic sequences. AAA81453 to AAA82414

XX CC represent specifically claimed *Neisseria meningitidis* genomic DNA

XX CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

XX CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to

XX CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

XX CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81322 to

XX CC AAA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF

XX CC sequences, which are all used in the exemplification of the present

XX CC invention. The nucleic acid sequences, protein sequences, and antibodies

XX CC against them, can be used in the manufacture of a composition. The

XX CC composition can be used as a medicament (or in the manufacture of a

XX CC medicament) for treating, preventing or diagnosing infection due to

XX CC *Neisserial* bacteria. For example, some of the identified proteins could

XX CC be components of vaccines against *Meningococcus* B; against all serotypes;

XX CC and/or against all pathogenic *Neisseriae*. Identification of sequences

XX CC from the bacterium will also facilitate production of biological probes,

XX CC particularly organism-specific probes. Attempts to make efficacious

XX CC *Meningococcus* B vaccines have failed mainly due to antigen tolerance.

XX CC Multivalent vaccines have also been tried but none have successfully

XX CC overcome antigenic variability. The provision of further, complete

XX CC sequences may provide an opportunity to identify secreted or surface

XX CC exposed proteins that may be presumed targets for the immune system and

XX CC which are not antigenically variable or at least more conserved than

XX CC other more variable regions.

XX SQ Sequence 12077 BP; 2643 A; 3404 C; 3147 G; 2881 T; 2 other;

Query Match 84.7%; Score 226.2; DB 21; Length 12077;

Best Local Similarity 97.0%; Pred. No. 8.2e-66;

Matches 263; Conservative 0; Mismatches 3; Indels 5; Gaps 3;

QY 1 AATTATGAACACAGCATCATCGTTTCGGTGGCGTTCGTTGCGTTGGCATACAGGTTG 60

DB 11627 AATTATGAACACAGCGCATCATCGTTTCGGTGGCGTTCGTTGCGTTGGCATACGCA-GTTG 11685

QY 61 CGGCTCAATCAATATGTAACCGTTTCGACACAGAACTTCAGGAACGTCGCGGTTTCG 120

DB 11686 CGGCTCAATCAATATGTAACCGTTTCGACTAGAACTTCAGGAACGTCGCGGTTTCG 11745

QY 121 CTTGGGCGTCA---CGAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCAT 177

DB 11746 CTTGGGCGTCAAGCCAAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCAT 11805

QY 178 CAACCTTACCGCAACTGTGGTGAAGCGGTGA-CCAATGCTATGTTACCAAGTGAATCAG 236

DB 11806 CAACCTTACCGCAACTGTGGTGAAGCGGTGAAGCAATGCTATGTTACCAAGTGAATCAG 11865

QY 237 CACAATCGGCTTACCACTTCGATGCAATT 267

DB 11866 CACAATCGGCTTACCACTTCGATGCAATT 11896

RESULT 9

ABZ38335/c

ID ABZ38335 standard; DNA; 420 BP.

XX AC ABZ38335;

XX DT 07-MAR-2003 (first entry)

XX DE *N. gonorrhoeae* nucleotide sequence SEQ ID 1259.

XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200279243-A2.

XX PT 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB02069.

XX PR 12-FEB-2001; 2001GB-0003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Masignani V, Monaci E;

XX WP1; 2003-058415/05.

XX P-PSDB; ABP77365.

XX DR New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a

XX PT medicament for treating or preventing *N. gonorrhoeae* infection

XX PS Disclosure; Page 279; 815pp; English.

XX CC The present invention relates to proteins from *Neisseria gonorrhoeae*.

XX CC Also disclosed are the nucleic acid molecules encoding the proteins and

XX CC antibodies that specifically bind to the proteins. The composition

XX CC comprising the protein, nucleic acid or antibody is useful for the

XX CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*

XX CC infection, this may be in the form of a vaccine or gene therapy.

XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid

XX CC molecules of the invention.

XX SQ Sequence 420 BP; 126 A; 78 C; 90 G; 126 T; 0 other;

Query Match 27.9%; Score 74.4; DB 25; Length 420;

Best Local Similarity 97.0%; Pred. No. 5.5e-15;

Matches 97; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 169 CATACGCATCAACTTACCGCAACTGTGGTAAGCGGTGA-CCAATGCTATGTTACGAG 227

DB 330 CACACGCATCAACTTACCGCAACTGTGGTAAGCGGTGAAGCGGTGCTATGTTACGAG 272

QY 228 TGTATACGCAACTCGCGTTACCACTTCGATGCAATT 267

DB 271 TGTATACGCAACTCGCGTTACCACTTCGATGCAATT 232

RESULT 10

ABZ38334

ID ABZ38334 standard; DNA; 144 BP.

XX AC ABZ38334;

XX DT 07-MAR-2003 (first entry)

XX DE *N. gonorrhoeae* nucleotide sequence SEQ ID 1257.

XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB02069.

XX PR 12-FEB-2001; 2001GB-0003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Masignani V, Monaci E;

XX WP1; 2003-058415/05.

XX P-PSDB; ABP77364.

XX PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a

XX PT medicament for treating or preventing *N. gonorrhoeae* infection

CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/3'-end sequence is selected from those defined in
CC	the specification. The primer sets can be used in antisense therapy and
CC	in gene therapy. The primers are useful for synthesising polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC	represent oligonucleotides, all of which are used in the exemplification
XX	of the present invention.
XX	
SQ	Sequence 827 BP; 200 A; 207 C; 234 G; 183 T; 3 other;
	Query Match 13.0%; Score 34.6; DB 22; Length 827;
	Best Local Similarity 55.4%; Pred. No. 0.23;
Matches	67; Conservative 0; Mismatches 54; Indels 0; Gaps 0
OY	100 TCAGAACGTGCCCGGTTTGGCTTGCGGCATCAATGCGGTAAAATCAGCAACCGCAG 159
Dd	383 TAAGACCGCGCGCTCTCACGCTGGCGTCACACACCGTGGCAGCTCATCTCTCG 422
OY	160 CAATCAAGGCATACGCATCACTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
Dd	423 CGAGACGACTGGCGGTCCACATCAACGAGGCCCGGAGATCGGCGTGGCCAGCACAA 483
OY	220 G 220
Dd	483 G 483
RESULT 12	
ID	ABX63506
AC	ABX63506 standard; cDNA; 1651 BP.
XN	
AX	ABX63506;
DT	26-FEB-2003 (first entry)
XX	
DE	Human cDNA #506 differentially expressed in activated vascular tissue.
KW	Human; Gene; ss; vascular tissue; cytostatic; atherosclerosis;
KW	cardiant; hypotensive; antidiabetic; gynaecological; vasotropic;
KW	cerebroprotective; gene therapy; vascular disease; cancer; coronary;
KW	artery disease; hypertension; diabetes; pre-eclampsia; restenosis;
KW	ischaemia-reperfusion injury; stroke;
OS	Homo sapiens.
XX	
US	US2002137081-A1.
PN	
PD	26-SEP-2002.
XX	
PP	08-JAN-2002; 2002US-0044030.
XX	
PR	28-JUL-2000; 2000US-222469P.
PR	08-JAN-2001; 2001US-260483P.
XX	
PA	(BAND/) BANDMAN O.
XX	
PI	Bandman O;
XX	
DR	WI; 2003-110597/10.
XX	
PPT	Combination for diagnosing, staging, treating, or monitoring the
PPT	progression of treatment of a vascular disease, e.g. atherosclerosis,
PPT	comprises several cDNAs that are differentially expressed in activated
PPT	vascular tissue -
XX	

Claim 1; Page -: 18pp; English.

PS This invention relates to a combination comprising several cDNAs that
XX are differentially expressed in activated vascular tissue. The invention
CC also discloses a high throughput method for detecting differentially
CC expressed cDNAs in a sample. The cDNAs of the invention may have
CC antiatherosclerotic; cytostatic; cardiac; hypotensive; antidiabetic;
CC gynaecological; vasotropic and cerebroprotective activities and may be
CC used in gene therapy. The cDNAs of the invention may be used in a
CC high-throughput methods for detecting differential expression of one or
CC more cDNAs in a sample, or screening several molecules or compounds to
CC identify a molecule or compound that specifically binds a cDNA of the
CC invention. A protein encoded by the cDNA may be used to screen several
CC molecules or compounds to identify a ligand that specifically binds to
CC the protein, or to produce or purify an antibody to the protein that can
CC be used to detect a protein in a sample or purify a natural or
CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of
CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary
CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-
CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used
CC for large-scale genetic or gene expression analysis of several new
CC nucleic acid molecules. Antibodies to the proteins encoded by the
CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic
CC or acute diseases associated with abnormalities in the expression,
CC amount or distribution of the protein. The present sequence
CC represents a cDNA of the invention that is differentially expressed in
CC activated vascular tissue.

CC Note: The sequence data for this patent did not form part of the
CC specification, but was obtained in electronic format directly from USPTO
CC at <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>.

XX Sequence 1651 BP; 386 A; 441 C; 410 G; 414 T; 0 other;

Query Match 13.0%; Score 34.6; DB 25; Length 1651;
Best Local Similarity 55.4%; Pred. No. 0.3;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGACGTCGCGGTTGCGTTCACCGGCGTCCACCAATGCCGTAAATCAGCAACCGCAG 159
Db 207 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACCAACACCGTGGCGAGCTCTCTCTCG 266.

QY 160 CAATCAAGGATAGCATCAACTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTAT 219

Db 267 CGAGACCGACTGCGGCTCCACATCAACGAGGCGCGGAGATCGGCGTGGCCAGCACCAA 326

QY 220 G 220

Db 327 G 327

RESULT 13

AAV18994

ID AAV18994 standard; cDNA; 1796 BP.

AC AAV18994;

DT 17-AUG-1998 (first entry)

XX Glutamine:fructose-6-phosphate amidotransferase TGC028-3 probe.

XX Glutamine:fructose-6-phosphate amidotransferase; GFAT; TGC028-3;

XX human; hypoglycaemia; diabetes; antidiabetic; therapy; probe; ss.

XX Homo sapiens.

XX EP824149-A2.

XX 18-FEB-1998.

XX 13-AUG-1997; 97EP-0113934.

XX 13-AUG-1996; 96JP-0213944.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hikichi Y, Nishi K, Shintani Y;

XX WPI; 1998-122309/12.

XX Glutamine:fructose-6-phosphate amidotransferase proteins - useful
PT for treating hypoglycaemia or to screen for antidiabetic agents

XX Example 3; Page 51-52; 77pp; English.

XX This sequence comprises a partial sequence of TGC028-3 (see AAV18988)
CC cDNA and codes for a portion of a novel human protein (see AAW37771)
CC having glutamine:fructose-6-phosphate amidotransferase (GFAT)
CC activity. It has been labelled with (alpha-32P)dCTP and used as a
CC probe to determine tissue specificity of gene expression by
CC Northern blotting. The invention relates to novel proteins (see
CC AAW29526, AAW37771-72) having GFAT activity, and DNAs encoding them (see
CC AAV18987-89), and their use for treating hypoglycaemia and for
CC identifying antidiabetic agents.

XX Sequence 1796 BP; 476 A; 451 C; 493 G; 376 T; 0 other;

Query Match 13.0%; Score 34.6; DB 19; Length 1796;

Best Local Similarity 55.4%; Pred. No. 0.31;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGACGTCGCGGTTGCGTTCACCGGCGTCCACCAATGCCGTAAATCAGCAACCGCAG 159

Db 1319 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACCAACACCGTGGCGAGCTCTCTCTCG 1378

QY 160 CAATCAAGGATAGCATCAACTTTACCGCAACTGTGGGTAAAGCGGTGACCAATGCTAT 219

Db 1379 CGAGACCGACTGCGGCTCCACATCAACGAGGCGCGGAGATCGGCGTGGCCAGCACCAA 1438

QY 220 G 220

Db 1439 G 1439

RESULT 14

AAV18988

ID AAV18988 standard; cDNA; 1845 BP.

AC AAV18988;

DT 17-AUG-1998 (first entry)

XX Glutamine:fructose-6-phosphate amidotransferase TGC028-3 cDNA.

XX Glutamine:fructose-6-phosphate amidotransferase; GFAT; TGC028-3;

XX human; hypoglycaemia; diabetes; antidiabetic; therapy; ds.

XX Homo sapiens.

XX EP824149-A2.

XX 18-FEB-1998.

XX 13-AUG-1997; 97EP-0113934.

XX 13-AUG-1996; 96JP-0213944.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hikichi Y, Nishi K, Shintani Y;

XX WPI; 1998-122309/12.

XX P-PSDB; AAW37771.

XX Glutamine:fructose-6-phosphate amidotransferase proteins - useful
PT for treating hypoglycaemia or to screen for antidiabetic agents

XX Claim 7; Page 43-44; 77pp; English.

PS This cDNA sequence codes for TGC028-3 (see AAW37771), a novel human

XX protein having glutamine:fructose-6-phosphate amidotransferase

XX (GFAT) activity and showing 74% homology to human GFAT. To isolate

CC TGC028-3 cDNA, a human brain cDNA library was probed with an

CC oligonucleotide (AAV18993), and DNA from a positive clone was

CC amplified using the same oligonucleotide as primer and introduced

CC into *Escherichia coli* DH10B. Transformants were screened by colony

CC PCR (see also AAV18991-92), and a colony which formed a 512-bp

CC amplified fragment was selected. The insert DNA of a plasmid from

CC this positive strain included the 1845 open reading frame of

CC TGC028-3. Plasmid pTBE1943 containing this DNA has been introduced

CC into *E. coli* DH10B and deposited as FERM BP-5622 (IPO 16004).

CC Also claimed are: a recombinant vector containing DNA (see also

CC AAV18987 and AAV18989) coding for a protein having GFAT activity (see

CC also AAW29526 and AAW37772) or its partial peptide, a transformant

CC transformed with the vector, a method for producing the protein by

CC culturing the transformant, an antibody against the protein, a

CC screening assay for compounds that inhibit the GFAT activity of the

CC protein, and a compound that inhibits GFAT activity. The GFAT, the

CC peptide fragment, or DNA coding for them, can be used in the

CC treatment and prophylaxis of hypoglycaemia. The GFAT inhibitor can

XX be used to treat diabetes.

XX Sequence 1845 BP; 486 A; 459 C; 505 G; 385 T; 0 other;

SQ

Query Match 13.0%; Score 34.6; DB 19; Length 1845;

Best Local Similarity 55.4%; Pred. No. 0.32;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAACGTGCGCGTTTGCTTGGCGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159

Db 1311 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACCAACCGTGGCGAGTCTCTCTCG 1370

QY 160 CAATGAAGCGATACGCATCACTTTACCGCACTGTGGTAAGCGGTGACCAATGCTAT 219

Db 1371 CGAGACCGACTGCGCGGTCCACATCAACGCGAGGCGGAGATCGCGTGGCCACCA 1430

QY 220 G 220

Db 1431 G 1431

RESULT 15

AAH14073

ID AAH14073 standard; cDNA; 1966 BP.

AC AAH14073;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:11221.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

FA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S; Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX

DR Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX Claim 8; SEQ ID 11221; 2537pp + CD ROM; English.

PS

XX The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 1966 BP; 471 A; 516 C; 472 G; 507 T; 0 other;

Query Match 13.0%; Score 34.6; DB 22; Length 1966;

Best Local Similarity 55.4%; Pred. No. 0.33;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAACGTGCGCGTTTGCTTGGCGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159

Db 363 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACCAACCGTGGCGAGTCTCTCTCG 422

QY 160 CAATGAAGCGATACGCATCACTTTACCGCACTGTGGTAAGCGGTGACCAATGCTAT 219

Db 423 CGAGACCGACTGCGCGGTCCACATCAACGCGAGGCGGAGATCGCGTGGCCACCA 482

QY 220 G 220

Db 483 G 483

Search completed: November 15, 2003, 00:35:42

Job time : 177.322 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1781.18 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-89

Perfect score: 267

Sequence: 1 AATATGACACACCATCA.....TTACACTTCGATGCAATT 267

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152338056 residues

Total number of hits satisfying chosen parameters: 45552784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hci.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.4	14.0	687	10	BG492320
2	35.2	13.2	580	13	BQ852510
3	34.6	13.0	350	9	AV659367
4	34.6	13.0	585	10	BG333201

RESULT 1
BG492320
LOCUS 602536235F1 NIH_MGC_59 Homo sapiens CDNA clone IMAGE:4555216 5',
DEFINITION 687 bp mRNA linear EST 27-MAR-2001
mRNA sequence.
ACCESSION BG492320
VERSION BG492320.1 GI:13453832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 687)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI444 row: e column: 09
High quality sequence stop: 548.

ALIGNMENTS

5	34.6	13.0	706	10	BG499592
6	34.6	13.0	827	9	AU124329
7	34.6	13.0	842	29	BZ729792
8	34.6	13.0	857	9	AU125108
9	34.6	13.0	911	13	BQ81308
10	33.8	12.7	811	28	BH422048
11	33.4	12.5	345	28	BH695730
12	33.4	12.5	661	12	B1165435
13	33.4	12.5	730	28	BH605016
14	33.4	12.5	883	10	BE113432
15	33.4	12.5	962	29	CNS02GGP
16	33	12.4	580	10	BF112720
17	33	12.4	939	13	EX430863
18	33	12.4	1041	12	BM423748
19	33	12.4	1201	13	EX385842
20	32.8	12.3	562	12	BI208441
21	32.8	12.3	624	12	BI930537
22	32.6	12.2	297	9	AV087447
23	32.6	12.2	443	12	BM030975
24	32.6	12.2	519	9	AW621988
25	32.6	12.2	1076	14	CB192570
26	32.4	12.1	733	28	BH140229
27	32.2	12.1	463	29	CNS013P2
28	32.2	12.1	746	12	BI154092
29	32.2	12.1	862	29	BZ787001
30	32.2	12.1	1201	13	EX324674
31	32	12.0	530	9	AW599664
32	32	12.0	803	28	BZ029668
33	32	12.0	804	28	BH536459
34	32	12.0	847	29	BZ961574
35	32	12.0	999	29	BZ493336
36	32	12.0	1021	29	CNS078BU
37	31.8	11.9	524	29	AG249081
38	31.8	11.9	589	12	BI167011
39	31.8	11.9	571	9	AA522229
40	31.8	11.9	681	28	BH527710
41	31.8	11.9	771	10	BF247178
42	31.8	11.9	810	28	BH481443
43	31.8	11.9	982	29	CNS0338Y
44	31.8	11.9	1201	13	EX438879
45	31.6	11.8	231	14	R29865

BG499592 602518007
AU124329 AU124329
BZ729792 QGFAT77TC
AU125108 AU125108
BQ81308 AGENCOURT
BH422048 BOGLS36TF
BH695730 BOMPE14TR
B1165435 RE05337.5
BH605016 BOHOAC2TF
BE113432 601148895
AL156306 Tetraodon
BF112720 EST440310
EX430863 EX430863
BM423748 AGENCOURT
EX385842 EX385842
BI208441 EST526481
BI930537 EST550426
AV087447 AV087447
BM030975 495552 WA
AW621988 EST312786
CB192570 AGENCOURT
BH140229 ZMMBB000
AL105728 Drosophila
BI154092 602870993
BZ787001 PUFHD79TB
EX324674 EX324674
AW599664 9491606.Y
BZ029668 oef79e06.
BH536459 BOHIQ56TR
BZ961574 PUDD822TD
BZ493336 BOKK575TR
AL433732 T3 end of
AG249081 Lotus japonicus
BI167011 RE07278.5
AA522229 V143C09.1
BH527710 BOGGL33TR
BF247178 601857849
BH481443 BOGRG76TF
AL225835 Tetraodon
EX438879 EX438879
R29865 12470 Lambd

```

FEATURES
  source
    Location/Qualifiers
      1..687
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4655216"
        /tissue_type="mucoepidermoid carcinoma"
        /lab_host="DH10B (TI phage-resistant)"
        /clone_lib="NH MGC 59"
        /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); SfiII (ggccctcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTGAGCCGAGCGGCACATG-dt(30)BN-3' (where B = A, C, G or A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
      BASE COUNT 197 a 166 c 146 g 178 t
      ORIGIN
        Query Match 14.0%; Score 37.4; DB 10; Length 687;
        Best Local Similarity 67.1%; Pred No. 0.52; Indels 0; Gaps 0;
        Matches 53; Conservative 0; Mismatches 26;
      QY 64 CTCATCAATATGTAACGGTTCGACACAGAACTTCAGGAACGTGCGCGTTTGCTT 123
      Db 561 CTCATTATGATCTAATTTAAACGAGGATAAATTTCAAACACAGTGTCTTCCTTCCCT 620
      QY 124 GGGCGTCACCATGCGTA 142
      Db 621 GGGCATCATCATTTCCATA 639
      RESULT 2
      LOCUS BQ852510 580 bp mRNA linear EST 14-AUG-2002
      DEFINITION QGB18D15.Y9.ab1 QG-ABCDI lettuce salinas Lactuca sativa CDNA clone
      ACCESSION QGB18D15, mRNA sequence.
      VERSION BQ852510.1 GI:22237975
      KEYWORDS EST.
      SOURCE Lactuca sativa
      ORGANISM Lactuca sativa
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
      asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
      Cichorieae; Lactuca.
      REFERENCE 1 (bases 1 to 580)
      AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
      Lin, H., van Damme, M., Lavelle, D., Cievalier, P., Ziegler, J., Ellison,
      P., Kolman, J., Slabaugh, K.S., Livingston, K., Zhou, Y., Lai, Z.,
      Church, S., Jackson, L. and Bradford, K.
      TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
      JOURNAL Unpublished
      COMMENT Contact: Alexander Kozik [R.W.Michelmore]
      Department of Vegetable Crops, R.W.Michelmore Lab
      University of California at Davis (UCD)
      Asmundson Hall, UCD, Davis, CA 95616, USA
      Tel: 1-(530)-742-1742
      Fax: 1-(530)-752-9659
      Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
      singleton, see http://cpdb.ucdavis.edu/ for details.
      Plate: QGB18 row; D column: 15.
      Location/Qualifiers
        1..580
          /organism="Lactuca sativa"
          /mol_type="mRNA"
          /cultivar="Salinas"
FEATURES
  source
    Location/Qualifiers
      1..350
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="GLCFWD02"
        /tissue_type="corresponding non cancerous liver tissue"
        /dev_stage="Adult"
        /lab_host="SCLR"
        /clone_lib="GLC"
        /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
      /db_xref="taxon:4236"
      /clone="QGB18D15"
      /lab_host="E.coli"
      /clone_lib="QG ABCDI lettuce salinas"
      /note="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cpdb.ucdavis.edu/ TAG_LIB-QG-ABCDI lettuce salinas TAG_TISSUE=Chemical induction TAG_SEQ=GTGATGCCGGG"
      BASE COUNT 183 a 102 c 143 g 152 t
      ORIGIN
        Query Match 13.2%; Score 35.2; DB 13; Length 580;
        Best Local Similarity 58.7%; Pred No. 2.4;
        Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
      QY 164 GAAGCATACGCATCAACTTTACCGCAACTGTGGTAAGCGGTGACCAATGCTATGTTA 223
      Db 244 GAAACATCGGATGATCTTGGTACCAATGTGGAGTTGGACTGCTACATGGAAGGAA 303
      QY 224 CCAGTGTATCAGCAATCGCGGTACCACTTCGATCGCAATT 267
      Db 304 CCAGTGGAGAATCAATCGGAATCAATTCGATGCGGTT 347
      RESULT 3
      LOCUS AV659367 350 bp mRNA linear EST 16-JAN-2002
      DEFINITION AV659367 GLC Homo sapiens cDNA clone GLCFWD02 3', mRNA sequence.
      ACCESSION AV659367
      VERSION AV659367.1 GI:9880381
      KEYWORDS EST.
      SOURCE Homo sapiens (human)
      ORGANISM Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      REFERENCE 1 (bases 1 to 350)
      AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
      Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
      Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
      Hu, G., Gu, J., Chen, Z. and Han, Z.
      TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
      JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
      MEDLINE 21625106
      PUBMED 11752456
      COMMENT Contact: Zequang Han
      Chinese National Human Genome Center at Shanghai
      351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
      201203, P. R. China
      Tel: 86-21-50801919 (ex.45)
      Fax: 86-21-50801922
      Email: hanzg@chgc.sh.cn
      This clone is available at CHGC in Shanghai.
      Location/Qualifiers
        1..350
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="GLCFWD02"
          /tissue_type="corresponding non cancerous liver tissue"
          /dev_stage="Adult"
          /lab_host="SCLR"
          /clone_lib="GLC"
          /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:

```

```

BASE COUNT      66 a      105 c      115 g      64 t
ORIGIN
Query Match      13.0%; Score 34.6; DB 9; Length 350;
Best Local Similarity 55.4%; Pred. No. 3;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAACGTGCGCGCTTTCCTTGGCGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 89 TAAGGACCGCGCGCTCTCACCGTGGCGTCCACCAACACCGTGGCGAGCTCCATCTCTCG 148
QY 160 CAATGAAGGCATACGCATCAACTTTACCGCAACTGTGGTAAGCGCGTGCACCAATGCTAT 219
DB 149 CGAGACCGACTGCGCGCTCCACATCAACGACGCGCGGAGATCGCGCTGCCGAGCACCAA 208
QY 220 G 220
DB 209 G 209

RESULT 4
BG333201      585 bp      mRNA      linear      EST 27-FEB-2001
LOCUS 602430915F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4548674 5',
DEFINITION mRNA sequence.
ACCESSION BG333201
VERSION BG333201.1 GI:13139639
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 585)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCMI238 row: n column: 03
High quality sequence stop: 585.
Location/Qualifiers
1. 585
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4548674"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 18"
/notes="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      138 a      148 c      165 g      134 t
ORIGIN
Query Match      13.0%; Score 34.6; DB 10; Length 585;
Best Local Similarity 55.4%; Pred. No. 3.8;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAACGTGCGCGCTTTCCTTGGCGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159

```

```

DB 375 TAAGGACCGCGCGCTCTCACCGTGGCGTCCACCAACACCGTGGCGAGCTCCATCTCTCG 434
QY 160 CAATGAAGGCATACGCATCAACTTTACCGCAACTGTGGTAAGCGCGTGCACCAATGCTAT 219
DB 435 CGAGACCGACTGCGCGCTCCACATCAACGACGCGCGGAGATCGCGCTGCCGAGCACCAA 494
QY 220 G 220
DB 495 G 495

RESULT 5
BG489592      706 bp      mRNA      linear      EST 27-MAR-2001
LOCUS 602518007F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4636608 5',
DEFINITION mRNA sequence.
ACCESSION BG489592
VERSION BG489592.1 GI:13451102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCMI395 row: n column: 01
High quality sequence stop: 650.
Location/Qualifiers
1. 706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4636608"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 18"
/notes="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      167 a      184 c      204 g      151 t
ORIGIN
Query Match      13.0%; Score 34.6; DB 10; Length 706;
Best Local Similarity 55.4%; Pred. No. 4.2;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAACGTGCGCGCTTTCCTTGGCGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 222 TAAGGACCGCGCGCTCTCACCGTGGCGTCCACCAACACCGTGGCGAGCTCCATCTCTCG 281
QY 160 CAATGAAGGCATACGCATCAACTTTACCGCAACTGTGGTAAGCGCGTGCACCAATGCTAT 219
DB 282 CGAGACCGACTGCGCGCTCCACATCAACGACGCGCGGAGATCGCGTGGCGAGCACCAA 341
QY 220 G 220

```

```

Db          342 G 342

RESULT 6
LOCUS      AUI24329
DEFINITION AUI24329 NT2RM2 Homo sapiens cDNA clone NT2RM2002030 5', mRNA
sequence.
ACCESSION AUI24329
VERSION    AUI24329.1 GI:10949045
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE 1 (bases 1 to 827)
AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
,Y., Sugano,S., Isogai,T.)
JOURNAL   Unpublished
COMMENT   Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute

FEATURES   source
            1..827
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="NT2RM2002030"
            /cell_type="teratocarcinoma"
            /cell_line="NT2"
            /clone_lib="NT2RM2"
            /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
            precursor cells"
BASE COUNT 200 a 207 c 234 g 183 t 3 others
ORIGIN

Query Match 13.0%; Score 34.6; DB 9; Length 827;
Best Local Similarity 55.4%; Pred. No. 4.5; Mismatches 0; Indels 0; Gaps 0;
Matches 67; Conservative 0;

QY 100 TCAGGAACGTGCGGGTTTCCTTGGGGTCCACCAATCCGCTAAATAATCAGCAACCGCAG 159
Db 363 TAAGGACCGCGCGCTCTCACCGTGGCGTCACCAACACCGTGGCGAGCTCCATCTCTCG 422
QY 160 CAATGAGGCATACGCATCACTTTACCGCACTGTGGTAAGCCGTCGACCAATGCTAT 219
Db 423 CGAGACCGACTCGCGGCTCCACATCAACGCGGCGGAGATCGCGTGGCGACGACCA 482
QY 220 G 220
Db 483 G 483

RESULT 7
LOCUS      BZ729792
DEFINITION OGFAT77TC ZM 0.7 1.5 kb Zea mays genomic clone ZMMEMa0241N10,
genomic survey sequence.
ACCESSION BZ729792
VERSION    BZ729792.1 GI:28703040
KEYWORDS   GSS.

SOURCE     Zea mays
ORGANISM   Zea mays

REFERENCE 1 (bases 1 to 842)
AUTHORS   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
TITLE     Zea mays
JOURNAL   Unpublished
COMMENT   Contact: Cathy Whitelaw
          Other_GSSs: OGFAT77TM
          TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TP
          Class: sheared ends.
FEATURES   Location/Qualifiers
            1..842
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMEMa0241N10"
            /clone_lib="ZM 0.7 1.5 kb"
            /note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
BASE COUNT 191 a 193 c 213 g 245 t
ORIGIN

Query Match 13.0%; Score 34.6; DB 29; Length 842;
Best Local Similarity 49.7%; Pred. No. 4.5; Mismatches 89; Indels 0; Gaps 0;
Matches 98; Conservative 0;

QY 18 TCATCGTTTCGGTGGTTCGTTGGTTGGTTCGATTAGCGTTGGCGCTCAATCAATG 77
Db 74 TCAACCCCTTCCCTATCTCGGACACTTAGCGGATGAGCCTTCCTTTAATCAATCGG 133
QY 78 TAACCGTTTCCGACCAAGAACTTCAGGAACGTGCGCGGTTTCCTTGGCGGTCACCAATG 137
Db 134 TCATTAAGCCCGGAGGACCCACACACACTAGTGTTTCTTGCTAGCTTTACAAAGC 193
QY 138 CGGTAAATATCAGCAACCCGACGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 194
Db 194 ACTTAAGAATAAGATGGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 250

RESULT 8
LOCUS      AUI25108
DEFINITION AUI25108 NT2RM4 Homo sapiens cDNA clone NT2RM4001049 5', mRNA
sequence.
ACCESSION AUI25108
VERSION    AUI25108.1 GI:10949824
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE 1 (bases 1 to 857)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
JOURNAL   HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
,Y., Sugano,S., Isogai,T.)
COMMENT   Unpublished
          Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute

SOURCE     Zea mays
ORGANISM   Zea mays

REFERENCE 1 (bases 1 to 842)
AUTHORS   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
TITLE     Zea mays
JOURNAL   Unpublished
COMMENT   Contact: Cathy Whitelaw
          Other_GSSs: OGFAT77TM
          TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TP
          Class: sheared ends.
FEATURES   Location/Qualifiers
            1..842
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMEMa0241N10"
            /clone_lib="ZM 0.7 1.5 kb"
            /note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
BASE COUNT 191 a 193 c 213 g 245 t
ORIGIN

Query Match 13.0%; Score 34.6; DB 29; Length 842;
Best Local Similarity 49.7%; Pred. No. 4.5; Mismatches 89; Indels 0; Gaps 0;
Matches 98; Conservative 0;

QY 18 TCATCGTTTCGGTGGTTCGTTGGTTGGTTCGATTAGCGTTGGCGCTCAATCAATG 77
Db 74 TCAACCCCTTCCCTATCTCGGACACTTAGCGGATGAGCCTTCCTTTAATCAATCGG 133
QY 78 TAACCGTTTCCGACCAAGAACTTCAGGAACGTGCGCGGTTTCCTTGGCGGTCACCAATG 137
Db 134 TCATTAAGCCCGGAGGACCCACACACACTAGTGTTTCTTGCTAGCTTTACAAAGC 193
QY 138 CGGTAAATATCAGCAACCCGACGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 194
Db 194 ACTTAAGAATAAGATGGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 250

```

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@nih.gov

HRI human cDNA project; 5'- and 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

Location/Qualifiers

1..857

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RM4001049"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_lib="NT2RM4"

/note="Vector: pME189FL3; mRNA from uninduced NT2 neuronal precursor cells"

BASE COUNT 201 a 217 c 238 g 196 t 5 others

ORIGIN

Query Match 13.0%; Score 34.6; DB 9; Length 857;

Best Local Similarity 55.4%; Pred. No. 4.5; Indels 0; Gaps 0;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGAACGTCGGCGTTCCTTGGCGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159

Db 349 TAAGGACCGGGCGCTCTCACCGTGGGCGTCACCAACCGTGGCGAGCTCCATCTCTCG 408

QY 160 CAATGAAGGATACGATCACTTACCGCACTGTGGTAAAGCGGTGACCAATGCTAT 219

Db 409 CGAGACCGACTGCGCGGTCCACATCAACGAGGCGCGAGATCGCGTGGCCAGCACCA 468

QY 220 G 220

Db 469 G 469

RESULT 9

BQ691308

LOCUS

DEFINITION AGNCOURT_8345892 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6250859

5' mRNA sequence. 911 bp mRNA linear EST 15-JUL-2002

ACCSSION BQ691308

VERSION BQ691308.1 GI:21816624

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: sgapbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LUCM394 row: j column: 12

High quality sequence stop: 677.

Location/Qualifiers

1..911

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6250859"

/tissue_type="ductal carcinoma, cell line"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

0;

0;

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_110"

/note="Organ: pancreas; Vector: pOTS7; Site: 1: XhoI;

Site: 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

BASE COUNT 216 a 234 c 255 g 201 t

5 others

Query Match 13.0%; Score 34.6; DB 13; Length 911;

Best Local Similarity 55.4%; Pred. No. 4.7;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGAACGTCGGCGTTCCTTGGCGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159

Db 403 TAAGGACCGGGCGCTCTCACCGTGGGCGTCACCAACCGTGGCGAGCTCCATCTCTCG 462

QY 160 CAATGAAGGATACGATCACTTACCGCACTGTGGTAAAGCGGTGACCAATGCTAT 219

Db 463 CGAGACCGACTGCGCGGTCCACATCAACGAGGCGCGAGATCGCGTGGCCAGCACCA 522

QY 220 G 220

Db 523 G 523

RESULT 10

BH422048/c

LOCUS

DEFINITION

BH422048.1

GI:17607776

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 811)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSSs: BQGLS36TR

Contact: Chris Town

TIGR Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..811

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="T01000DH3"

/db_xref="taxon:3712"

/clone="BQGLS36"

/clone_lib="BQGL"

/note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHS01 using BstXI linkers"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

0;

0;

```

LOCUS
DEFINITION 6 TGAACACACCATCATCGTTTCGGCTGCGTTCGTTGGTGGCATTAGCAGGTGCGGCT 65
Drosophila melanogaster cDNA clone RE05337 5, mRNA sequence.
ACCESSION BI165435
VERSION BI165435.1 GI:14631241
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (Bases 1 to 861)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phuanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic:AB003518; arm:3L [18805396,19084925]
estimated-cyto:75E4-76A1: 04/11/2001
Plate: RE.53 row: D column: 1
High quality sequence stop: 561.
Location/Qualifiers
1..661
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE05337"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/Note="Organ: embryo; Vector: pFlc1; Site:1; XhoI; Site:2;
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 242 a 141 c 138 g 140 t
ORIGIN
Query Match 12.5%; Score 33.4; DB 12; Length 661;
Best Local Similarity 49.2%; Pred. No. 9.8;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 77 GTAACGTTTCGACCAAACTTCAGAACTGCCGCTTGCCTGGCGTCCACCAAT 136
Db 2 GTTACCGTTTCGGTCGCTAACGACGAGCAAAATCGAGAGCTGGCTTCATTAA 61
QY 137 GCGCTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAACTTTACCGCAACTGTG 196
Db 62 GCCAAACAATCAAAACAACAAGACCGAATATATGAAGTAACAGTTCAGAAAACAGAC 121
QY 197 GTTAAGCGGTGACCAATGCTATGTTACCATGTATATCAGCACAATCGCGGTTTACCCT 255
Db 122 GAAATTCTCCGCAACTCGCGGTAAAGTGGGAATAAAACAAATTGGATATACAATT 180
RESULT 13
BH605016
LOCUS
DEFINITION BOHO02TF BOHO Brassica oleracea genomic clone BOHO02, genomic
survey sequence.
ACCESSION BH605016
VERSION BH605016.1 GI:17857462
KEYWORDS GSS.

```

```

LOCUS
DEFINITION 6 TGAACACACCATCATCGTTTCGGCTGCGTTCGTTGGTGGCATTAGCAGGTGCGGCT 65
Drosophila melanogaster cDNA clone RE05337 5, mRNA sequence.
ACCESSION BI165435
VERSION BI165435.1 GI:14631241
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (Bases 1 to 861)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phuanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic:AB003518; arm:3L [18805396,19084925]
estimated-cyto:75E4-76A1: 04/11/2001
Plate: RE.53 row: D column: 1
High quality sequence stop: 561.
Location/Qualifiers
1..661
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE05337"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/Note="Organ: embryo; Vector: pFlc1; Site:1; XhoI; Site:2;
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 242 a 141 c 138 g 140 t
ORIGIN
Query Match 12.5%; Score 33.4; DB 12; Length 661;
Best Local Similarity 49.2%; Pred. No. 9.8;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 77 GTAACGTTTCGACCAAACTTCAGAACTGCCGCTTGCCTGGCGTCCACCAAT 136
Db 2 GTTACCGTTTCGGTCGCTAACGACGAGCAAAATCGAGAGCTGGCTTCATTAA 61
QY 137 GCGCTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAACTTTACCGCAACTGTG 196
Db 62 GCCAAACAATCAAAACAACAAGACCGAATATATGAAGTAACAGTTCAGAAAACAGAC 121
QY 197 GTTAAGCGGTGACCAATGCTATGTTACCATGTATATCAGCACAATCGCGGTTTACCCT 255
Db 122 GAAATTCTCCGCAACTCGCGGTAAAGTGGGAATAAAACAAATTGGATATACAATT 180
RESULT 13
BH605016
LOCUS
DEFINITION BOHO02TF BOHO Brassica oleracea genomic clone BOHO02, genomic
survey sequence.
ACCESSION BH605016
VERSION BH605016.1 GI:17857462
KEYWORDS GSS.

```

```

LOCUS
DEFINITION 6 TGAACACACCATCATCGTTTCGGCTGCGTTCGTTGGTGGCATTAGCAGGTGCGGCT 65
Drosophila melanogaster cDNA clone RE05337 5, mRNA sequence.
ACCESSION BI165435
VERSION BI165435.1 GI:14631241
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (Bases 1 to 861)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phuanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic:AB003518; arm:3L [18805396,19084925]
estimated-cyto:75E4-76A1: 04/11/2001
Plate: RE.53 row: D column: 1
High quality sequence stop: 561.
Location/Qualifiers
1..661
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE05337"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/Note="Organ: embryo; Vector: pFlc1; Site:1; XhoI; Site:2;
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 242 a 141 c 138 g 140 t
ORIGIN
Query Match 12.5%; Score 33.4; DB 12; Length 661;
Best Local Similarity 49.2%; Pred. No. 9.8;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 77 GTAACGTTTCGACCAAACTTCAGAACTGCCGCTTGCCTGGCGTCCACCAAT 136
Db 2 GTTACCGTTTCGGTCGCTAACGACGAGCAAAATCGAGAGCTGGCTTCATTAA 61
QY 137 GCGCTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAACTTTACCGCAACTGTG 196
Db 62 GCCAAACAATCAAAACAACAAGACCGAATATATGAAGTAACAGTTCAGAAAACAGAC 121
QY 197 GTTAAGCGGTGACCAATGCTATGTTACCATGTATATCAGCACAATCGCGGTTTACCCT 255
Db 122 GAAATTCTCCGCAACTCGCGGTAAAGTGGGAATAAAACAAATTGGATATACAATT 180
RESULT 13
BH605016
LOCUS
DEFINITION BOHO02TF BOHO Brassica oleracea genomic clone BOHO02, genomic
survey sequence.
ACCESSION BH605016
VERSION BH605016.1 GI:17857462
KEYWORDS GSS.

```

SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 730)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BOHOA02TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: IF
Class: sheared ends.
Location/Qualifiers
1..730
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="Tol000DH3"
/db_xref="taxon:3712"
/clone="BOHOA02"
/clone_lib="BOHO"
/note="Vector: pHOs1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOs1 using BstXI linkers"
227 a 213 c 164 g 126 t

Query Match 12.5%; Score 33.4; DB 28; Length 730;
Best Local Similarity 53.4%; Pred. No. 10;
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 56 GTTCCGCTCAATCAATGTAACCGTTCCGACGAGAACTTCAGCACTGCCGCG 115
Db GTTCACCGAGAACCAAGAGCTCCATTACAAACAGATCGCCAGCGTCAGACTGCA 591
QY 116 TTTCGCTTGGCGTCACCAATGCCGTAATAATCAGCAACCGCAATGAAGGCATACGC 175
Db TCACGCTGGAGTTAAATTCGCGGAATATGCAGGAACACAGAACCAAGGCCACGT 651
QY 176 ATCACTTTAC 186
Db ACGAGCTTCC 662

RESULT 14
LOCUS BE313432/c
DEFINITION BE313432 883 bp mRNA linear EST 26-OCT-2000
60114895F1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:3501886 5',
mRNA sequence.
ACCESSION BE313432
VERSION BE313432
KEYWORDS BE313432.1 GI:9133328
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 883)
TITLE NIH-MGC http://mgi.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCMI72 row: m column: 23
High quality sequence stop: 222.
Location/Qualifiers
1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3501886"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
188 a 277 c 281 g 137 t

BASE COUNT
ORIGIN

Query Match 12.5%; Score 33.4; DB 10; Length 883;
Best Local Similarity 62.7%; Pred. No. 11;
Matches 52; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 185 ACCGCAACTGTGGTAAGCGCGTGACCAATGCTATGTTACCACTGTAATCAGCACATCG 244
Db ACCCAAGCGCCCTAAGGCATCAGCACCCCATGATACCCGCCCATCAGCACATTCG 269
QY 245 GCGTTACCACTTCCGATGCAATT 267
Db GCCATTTCACACCAATGCAACT 246

RESULT 15
LOCUS CNS02GGP/c
DEFINITION CNS02GGP 962 bp DNA linear GSS 01-SEP-2000
137J21 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL196306
KEYWORDS ALL96306.1 GI:7834456
SOURCE GSS; genome survey sequence.
ORGANISM Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
Roest Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645
2
Roest Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10899143
3 (bases 1 to 962)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 41.5874 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-89

Perfect score: 267

Sequence: 1 AATTATGACACACGATCA.....TTACACTTCGATGCAATT 267

Scoring table: IDENTITY NJC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34.6	13.0	1796	2	US-08-911-445-17
2	34.6	13.0	1796	3	US-09-182-983-17
3	34.6	13.0	1845	2	US-08-911-445-5
4	34.6	13.0	1845	3	US-09-182-983-5
5	34.6	13.0	2046	2	US-08-911-445-6
6	34.6	13.0	2046	3	US-09-182-983-6
7	30.4	11.4	1466	4	US-08-620-332D-914
8	29.4	11.0	1455	3	US-08-276-531-33
9	29	10.9	3947	3	US-08-975-762-47
10	29	10.9	3947	3	US-08-295-028-47
11	29	10.9	3947	4	US-09-106-582-47
12	28.2	10.6	1665	3	US-09-292-768-69
13	28.2	10.6	1764	3	US-08-289-532-17
14	28	10.5	8040	1	US-08-536-231-1
15	28	10.5	8040	3	US-09-100-804-1
16	28	10.5	8043	5	PCT-US94-09943-1
17	28	10.5	8119	3	US-08-290-640-45
18	28	10.5	4403765	3	US-09-103-840A-2
19	28	10.5	4411529	3	US-09-103-840A-1
20	27.8	10.4	2430	1	US-08-062-368-1
21	27.8	10.4	1796	4	US-08-647-452A-1
22	27.6	10.3	1103	1	US-08-181-271A-3
23	27.6	10.3	1103	1	US-08-449-315-3
24	27.6	10.3	1103	1	US-08-444-803-3
25	27.6	10.3	1103	1	US-08-449-043-3
26	27.6	10.3	1103	1	US-08-456-265A-3
27	27.6	10.3	1103	1	US-08-455-416-3

28	27.6	10.3	1103	1	US-08-455-244-3	Sequence 3, Appl
29	27.6	10.3	1103	1	US-08-454-876-3	Sequence 3, Appl
30	27.6	10.3	1103	2	US-08-457-364-3	Sequence 3, Appl
31	27.6	10.3	1103	2	US-08-456-262-3	Sequence 3, Appl
32	27.6	10.3	1103	2	US-08-456-240-3	Sequence 3, Appl
33	27.6	10.3	1103	2	US-08-455-736-3	Sequence 3, Appl
34	27.6	10.3	1103	2	US-08-971-217-3	Sequence 3, Appl
35	27.6	10.3	1103	3	US-08-329-799-38	Sequence 3, Appl
36	27.6	10.3	1103	3	US-09-350-600-3	Sequence 3, Appl
37	27.6	10.3	1224	4	US-09-615-192A-363	Sequence 3, Appl
38	27.6	10.3	2200	4	US-09-504-358-27	Sequence 27, Appl
39	27.6	10.3	2200	4	US-09-504-314-27	Sequence 27, Appl
40	27.6	10.3	11471	4	US-09-504-358-16	Sequence 16, Appl
41	27.6	10.3	11471	4	US-09-504-314-16	Sequence 16, Appl
42	27.6	10.3	12124	1	US-08-181-271A-36	Sequence 36, Appl
43	27.6	10.3	12124	1	US-08-449-315-36	Sequence 36, Appl
44	27.6	10.3	12124	1	US-08-444-803-36	Sequence 36, Appl
45	27.6	10.3	12124	1	US-08-449-043-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-08-911-445-17
; Sequence 17, Application US/08911445
; Patent No. 5876713
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,445
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 213944-1996
; FILING DATE: 13-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-911-445-17

Query Match 13.0%; Score 34.6; DB 2; Length 1796;
Best Local Similarity 55.4%; Pred. No. 0.043;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGGAACGTGGCGTTTCCTTGGCGTCCATGCGTAAAAATCAGAACCCGAG 159
DB 1319 TAAGACCGCGCTCTCACCGTGGCGTCCAAACCGGTGGCGTCTCTCTCG 1378

QY 160 CAATGAAGGCATACGCATCACTTTACCGCAACTGTGGTAAGCGGTGACCAATGCTAT 219
DB 1379 CGAGACCGACTGCGCGGTCTCCACATCAACGAGGCGGAGATCGCGTGGCCAGACCAA 1438
QY 220 G 220
DB 1439 G 1439

RESULT 2
US-09-182-983-17
; Sequence 17, Application US/09182983
; Patent No. 6207431
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09182,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-182-983-17

Query Match 13.0%; Score 34.6; DB 3; Length 1796;
Best Local Similarity 55.4%; Pred. No. 0.043; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGGAAGTCGCGCGTTGCTTGGCGGTACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 1319 TAAGGACCGCGCGGTCTCCACGTTGGCGGTACCAACACCGGTGGCGAGTCCATCTCTCG 1378
QY 160 CAATGAAGGCATACGCATCACTTTACCGCAACTGTGGTAAGCGGTGACCAATGCTAT 219
DB 1379 CGAGACCGACTGCGCGGTCTCCACATCAACGAGGCGGAGATCGCGTGGCCAGACCAA 1438
QY 220 G 220
DB 1439 G 1439

RESULT 3
US-08-911-445-5
; Sequence 5, Application US/08911445

; Patent No. 5876713
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,445
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 213944-1996
; FILING DATE: 13-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-911-445-5

Query Match 13.0%; Score 34.6; DB 2; Length 1845;
Best Local Similarity 55.4%; Pred. No. 0.043; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGGAAGTCGCGCGTTGCTTGGCGGTACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 1311 TAAGGACCGCGCGGTCTCCACGTTGGCGGTACCAACACCGGTGGCGAGTCCATCTCTCG 1370
QY 160 CAATGAAGGCATACGCATCACTTTACCGCAACTGTGGTAAGCGGTGACCAATGCTAT 219
DB 1371 CGAGACCGACTGCGCGGTCTCCACATCAACGAGGCGGAGATCGCGTGGCCAGACCAA 1430
QY 220 G 220
DB 1431 G 1431

RESULT 4
US-09-182-983-5
; Sequence 5, Application US/09182983
; Patent No. 6207431
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-5
;
; Query Match 13.0%; Score 34.6; DB 3; Length 1845;
; Best Local Similarity 55.4%; Pred. No. 0.043;
; Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
;
; QY 100 TCAGGAACGTGCGCGCTTTGCCCTTGGCGGTCCACCAATGCCGTAACCAATCAGCAACCGCAG 159
; DB 1311 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACACCGTGGCGCAGCTCACTCTCG 1370
;
; QY 160 CAATGAAGCATACGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGTAT 219
; DB 1371 CGAGACCGACTGCGGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACAA 1430
;
; QY 220 G 220
; DB 1431 G 1431
;
; RESULT 5
; US-09-11-445-6
; Sequence 6, Application US/08911445
; Patent No. 5876713
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,445
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 213944-1996
; FILING DATE: 13-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-6
;
; Query Match 13.0%; Score 34.6; DB 2; Length 2046;
; Best Local Similarity 55.4%; Pred. No. 0.046;
; Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
;
; QY 100 TCAGGAACGTGCGCGCTTTGCCCTTGGCGGTCCACCAATGCCGTAACCAATCAGCAACCGCAG 159
; DB 1311 TAAGGACCGCGCGCTCTCACCGTGGCGGTCCACCAACCGTGGCGCAGCTCACTCTCG 1370
;
; QY 160 CAATGAAGCATACGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGTAT 219
; DB 1371 CGAGACCGACTGCGGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACAA 1430
;
; QY 220 G 220
; DB 1431 G 1431
;
; RESULT 6
; US-09-182-983-6
; Sequence 6, Application US/09182983
; Patent No. 6207431
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-6

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-5
;
; Query Match 13.0%; Score 34.6; DB 3; Length 1845;
; Best Local Similarity 55.4%; Pred. No. 0.043;
; Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
;
; QY 100 TCAGGAACGTGCGCGCTTTGCGCTTCCACCAATGCCGTAACCAATCAGCAACCGCAG 159
; DB 1311 TAAGGACCGCGCGCTCTCACCGTGGCGTCCACACCGTGGCGCAGCTCACTCTCG 1370
;
; QY 160 CAATGAAGCATACGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
; DB 1371 CGAGACCGACTGCGGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACAA 1430
;
; QY 220 G 220
; DB 1431 G 1431
;
; RESULT 5
; US-09-11-445-6
; Sequence 6, Application US/08911445
; Patent No. 5876713
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,445
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 213944-1996
; FILING DATE: 13-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-6
;
; Query Match 13.0%; Score 34.6; DB 2; Length 2046;
; Best Local Similarity 55.4%; Pred. No. 0.046;
; Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
;
; QY 100 TCAGGAACGTGCGCGCTTTGCGCTTCCACCAATGCCGTAACCAATCAGCAACCGCAG 159
; DB 1311 TAAGGACCGCGCGCTCTCACCGTGGCGTCCACACCGTGGCGCAGCTCACTCTCG 1370
;
; QY 160 CAATGAAGCATACGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
; DB 1371 CGAGACCGACTGCGGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACAA 1430
;
; QY 220 G 220
; DB 1431 G 1431
;
; RESULT 6
; US-09-182-983-6
; Sequence 6, Application US/09182983
; Patent No. 6207431
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; TITLE OF INVENTION: AMIDOTRANSFERASE, ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FETHERSTONHAUGH & Co.,
; REFERENCE/DOCKET NUMBER: 28605-30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613)-232-2486
; TELEFAX: (613)-232-8440
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-182-983-6

```

Query Match 13.0%; Score 34.6; DB 3; Length 2046;
 Best Local Similarity 55.4%; Pred. No. 0.046;
 Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGACGTCGCCGTTGGCGTCACCAATGCCGTAATAATCAGCAACCGCAG 159
 DB 1311 TAAGACCGCGCGCTCTCACCCTGGCGCTCACCACACCGTGGCGAGCTCCATCTCTCG 1370

QY 160 CAATGAAGGCATAGCATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
 DB 1371 CGAGACCGACTGCGCGCTCCACATCAACGACGCGGAGATCGGCGTGCCAGCACAA 1430

QY 220 G 220
 DB 1431 G 1431

RESULT 7

US-09-620-312D-914
 ; Sequence 914, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Dmanac, Radole T.

; TITLE OF INVENTION: No. 6569662a1 Nucleic Acids and

; FILE REFERENCE: Polypeptides

; CURRENT APPLICATION NUMBER: US/09/620.312D

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt FL_genes Version 1.0

; SEQ ID NO 914

; LENGTH: 1466

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (779)..(1213)

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(1466)

; OTHER INFORMATION: n = a,t,c or g

US-09-620-312D-914

Query Match 11.4%; Score 30.4; DB 4; Length 1466;
 Best Local Similarity 57.3%; Pred. No. 1.1;
 Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 89 GACGAGAACTCAGGACGTCGCCGTTTCCTTGGCGCTCACCACCAATGCCGTAATAATC 148

DB 478 GAGGAGAAGCTTCGGAACACGCTGCCCTCCCGCAAGGCGCTCAGTTTCCATTATGTC 537

QY 149 AGCAACCCGCAATGAAGGCATAGCATCAACTTT 184

DB 538 AACACACGCTTGAAGACATGAGGAGCAACACGTT 573

RESULT 8

US-09-276-531-33

; Sequence 33, Application US/09276531

; Patent No. 6183968

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Reddy, Roopa

; APPLICANT: Guegler, Karl J.

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING

; NUMBER OF SEQUENCES: 134

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/276.531

; FILING DATE: Herewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/079,677

; FILING DATE: March 27, 1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Lynn E. Murry, Ph.D.

; REGISTRATION NUMBER: 42,918

; REFERENCE/DOCKET NUMBER: PA-0008 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1455 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BRAITUT01

; CLONE: 746308

US-09-276-531-33

Query Match 11.0%; Score 29.4; DB 3; Length 1455;
 Best Local Similarity 56.8%; Pred. No. 2.4;
 Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 18 TCATCGTTTCGGCTCGGTTGCGTTGGCATTAGCAGGTTGCGGCTCAATCAATAATG 77

DB 94 TCGCATCTCTGCATCTCTGCGGCGGTGAGCAATCGAGGCGCTGGCTCAGTCAATAAG 153

QY 78 TAAACGTTTCCACCAGAACTTCAGGAACGTC 112

DB 154 GATCGGGGCCAACACGACATTCACAGTAGTTCC 188

RESULT 9

US-08-975-762-47/c

; Sequence 47, Application US/08975762

```
Patent No. 6207169
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA: Patent In Release #1.0, Version #1.30
APPLICATION NUMBER: US/08/975,762
FILING DATE: 21-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 3947 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-975-762-47

Query Match 10.9%; Score 29; DB 3; Length 3947;
Best Local Similarity 49.7%; Pred. No. 5.2;
Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 71 AATAATGTAACCGTTTCGACAGCAAACTTCAGGAACGTCGCGGTTTGCTTGGGGTGC 130
DB 2148 AATAGTCTGCTGATTTCTTATAAGTTAAATATAGCTCGTGGCGAATTCGGC 2089

QY 131 ACCAATGCGGTAAATAATCAGCAACCCGACGCAATGAAGGCATACGCATCAACTTTACCGCA 190
DB 2088 ACGAGTATTAGGGGCGAGAAATTTTCACAGCATTCGCAAACTGCAAAAGTCTTTACCGGA 2029

QY 191 ACTGTGGGTAAGCGGTGACCAATGCTAT 219
DB 2028 AGTTTGGCAGAGCAGCAGCAAGGTTAT 2000

RESULT 10
US-09-295-028-47/c
Sequence 47, Application US/09295028
Patent No. 6277381
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
AND TREATMENT OF EHRlichia INFECTION
FILE REFERENCE: 210121.439C4
CURRENT APPLICATION NUMBER: US/09/295,028
CURRENT FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 85
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 47
```

```
LENGTH: 3947
TYPE: DNA
ORGANISM: Ehrlichia sp.
US-09-295-028-47

Query Match 10.9%; Score 29; DB 3; Length 3947;
Best Local Similarity 49.7%; Pred. No. 5.2;
Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 71 AATAATGTAACCGTTTCGACAGCAAACTTCAGGAACGTCGCGGTTTGCTTGGGGTGC 130
DB 2148 AATAGTCTGCTGATTTCTTATAAGTTAAATATAGCTCGTGGCGAATTCGGC 2089

QY 131 ACCAATGCGGTAAATAATCAGCAACCCGACGCAATGAAGGCATACGCATCAACTTTACCGCA 190
DB 2088 ACGAGTATTAGGGGCGAGAAATTTTCACAGCATTCGCAAACTGCAAAAGTCTTTACCGGA 2029

QY 191 ACTGTGGGTAAGCGGTGACCAATGCTAT 219
DB 2028 AGTTTGGCAGAGCAGCAGCAAGGTTAT 2000

RESULT 11
US-09-106-582-47/c
Sequence 47, Application US/09106582
Patent No. 6306402
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
THERAPY
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,582
FILING DATE: 29-JUN-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 3947 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-106-582-47

Query Match 10.9%; Score 29; DB 4; Length 3947;
Best Local Similarity 49.7%; Pred. No. 5.2;
Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 71 AATAATGTAACCGTTTCGACAGCAAACTTCAGGAACGTCGCGGTTTGCTTGGGGTGC 130
DB 2148 AATAGTCTGCTGATTTCTTATAAGTTAAATATAGCTCGTGGCGAATTCGGC 2089

QY 131 ACCAATGCGGTAAATAATCAGCAACCCGACGCAATGAAGGCATACGCATCAACTTTACCGCA 190
```

Db 2088 ACAGTATTAGGGGAGAGAAATTTTACAGCATTCGCAAAATGCAACAAAGTCTTTACCGGA 2029
Qy 191 ACTGTGGGTAAAGCGGTCACCAATCCTAT 219
Db 2028 AGTTTGGCAAGAGCAGCAGCAAGGTTAT 2000

RESULT 12
US-09-292-768-69
; Sequence 69, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wslr13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase protein variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-69
Query Match 10.6%; Score 28.2; DB 3; Length 1665;
Best Local Similarity 47.9%; Pred. No. 6.6;
Matches 81; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 61 CGGCTCATCATATATGTAAACGGTTTCGACCAAGAACTTCAGGAACGTGCGCGTTTGC 120
Db 753 CGAGCCATCTCGAGGCGCATCGTGGAGCAGCAAGCTCAAGAGAGCGCGGAGTTTG 812
Qy 121 CTTGGCGGTCAACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGGCATACGCATCAA 180
Db 813 CGCGAGACATATTAGAGTACTCTTTAGATCGAGAGGATAGCCAGATCAAGTCCC 872
Qy 181 CTTTACCCCAACTGTGGGTAGCGCGTGAACCAATGCTATGTACCATGT 229
Db 873 CATCACCAACCAAGCCATCAAAAGCGCTTCATCTTCGACACGCTTCTCAGCG 921

RESULT 13
US-09-269-592-17/c
; Sequence 17, Application US/09269592
; Patent No. 6235500
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G.
; APPLICANT: Sanders, Kevin
; TITLE OF INVENTION: OXYGEN-BINDING HEME PROTEINS INCORPORATING
; FILE OF INVENTION: CIRCULARLY-PERMUTED GLOBINS
; FILE REFERENCE: 22010-140
; CURRENT APPLICATION NUMBER: US/09/269,592
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: PCT/US97/17294

; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 60/026,831
; PRIOR FILING DATE: 1996-09-26
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(1752)
US-09-269-592-17

Query Match 10.6%; Score 28.2; DB 3; Length 1764;
Best Local Similarity 57.3%; Pred. No. 6.8;
Matches 51; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 124 GGGCGTCACCAATGCGGTAAATAATCAGCAACCGCAGCAATGAAGCATACGCATCAACTT 183
Db 130 GAGCGTTCGGCATGTGTCACATGAGCAACAGCGTTAGTCAACCGTCAGCAACTTTT 71
Qy 184 TACCGCAACTGTGGTAAAGCGCGTGACCA 212
Db 70 TACCATGACCTTAACCTGAGCGGAACCA 42

RESULT 14
US-08-596-291-1
; Sequence 1, Application US/08596291
; Patent No. 5821075
; GENERAL INFORMATION:
; APPLICANT: GONEZ, LEONEL JORGE
; APPLICANT: SARAS, JAN
; APPLICANT: CLAESSON-WELSH, LENA
; APPLICANT: HELDIN, CARL-HENRIK
; TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
; TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,291
; FILING DATE: 03-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/115,573
; FILING DATE: 01-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: L0461/7000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; TELEX: 92-1742 EZEKIEL
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..7475
; US-08-596-291-1

Query Match      10.5%; Score 28; DB 1; Length 8040;
Best Local Similarity 56.5%; Pred. No. 16;
Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 15 GCATCATCGTTTCGGTCGCTGCGTTGGCATTAGCAGGTTGCGGCTCAATCAATA 74
DB 220 GCTTCATCATTTCTCCATGGTCTGCTGTTGTCGCCATCTGAGTGTGTCTATTACAG 279

QY 75 ATGTAACCGTTTCGACACGAGAACTTCAGGAA 106
DB 280 ATGAAATATTTCCAAATCAGATCTTCGAGCA 311

; TOPLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..7475
; US-09-100-804-1

Query Match      10.5%; Score 28; DB 3; Length 8040;
Best Local Similarity 56.5%; Pred. No. 16;
Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 15 GCATCATCGTTTCGGTCGCTGCGTTGGCATTAGCAGGTTGCGGCTCAATCAATA 74
DB 220 GCTTCATCATTTCTCCATGGTCTGCTGTTGTCGCCATCTGAGTGTGTCTATTACAG 279

QY 75 ATGTAACCGTTTCGACACGAGAACTTCAGGAA 106
DB 280 ATGAAATATTTCCAAATCAGATCTTCGAGCA 311
```

Search completed: November 15, 2003, 08:09:32
Job time : 46.5874 secs

```

; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..7475
; US-08-596-291-1

Query Match      10.5%; Score 28; DB 1; Length 8040;
Best Local Similarity 56.5%; Pred. No. 16;
Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 15 GCATCATCGTTTCGGTCGCTGCGTTGGCATTAGCAGGTTGCGGCTCAATCAATA 74
DB 220 GCTTCATCATTTCTCCATGGTCTGCTGTTGTCGCCATCTGAGTGTGTCTATTACAG 279

QY 75 ATGTAACCGTTTCGACACGAGAACTTCAGGAA 106
DB 280 ATGAAATATTTCCAAATCAGATCTTCGAGCA 311

RESULT 15
US-09-100-804-1
; Sequence 1, Application US/09100804
; Patent No. 6066472
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, LEONEL JORGE
; APPLICANT: SARAS, JAN
; APPLICANT: CLAESSON-WELSH, LENA
; APPLICANT: HELDIN, CARL-HENRIK
; TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
; TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.804
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/596.291
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/115.573
; FILING DATE: 01-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09943
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: L0461/7003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 151.197 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-89

Perfect score: 267
Sequence: 1 AATTATGACACACGCATCA.....TTACCACTTCGATGCAATT 267

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	267	10	US-09-928-457-89
2	34.6	13.0	1551	13	US-10-044-090-506
3	34.6	13.0	1796	9	US-09-771-838A-17
4	34.6	13.0	1845	9	US-09-771-838A-5
5	34.6	13.0	2046	9	US-09-771-838A-6
6	34.6	13.0	3023	9	US-09-079-892-4
7	34.6	13.0	3048	13	US-10-044-090-685
8	31.4	11.8	369	10	US-09-867-701-3070
9	31.2	11.7	1601042	12	US-10-027-632-59064
10	31.2	11.7	1601042	13	US-10-027-632-59064
11	30.4	11.4	1466	14	US-10-037-270-914
12	30	11.2	74868	14	US-10-175-523-67
13	29.4	11.0	1473	11	US-09-796-753-47
14	29.4	11.0	1481	14	US-10-198-846-11017
15	29.4	11.0	1741	13	US-10-098-841-67
16	29.4	11.0	2211	10	US-09-738-626-1688

29.4 11.0 5144 11 US-09-919-039-70 Sequence 70, Appl
C 17 29.4 11.0 3309400 10 US-09-738-626-1 Sequence 1, Appl
C 18 29.4 10.9 384 9 US-09-879-792-6 Sequence 6, Appl
C 19 29.2 10.9 384 10 US-09-867-701-2678 Sequence 2678, Ap
C 20 29.2 10.9 384 10 US-09-867-701-142 Sequence 142, App
C 21 29 10.9 333 10 US-09-867-701-142 Sequence 1, Appl
C 22 29 10.9 1100 9 US-09-859-822-1 Sequence 47, Appl
C 23 29 10.9 3947 9 US-09-159-469-47 Sequence 47, Appl
C 24 29 10.9 3947 9 US-09-798-042-47 Sequence 1370, Ap
C 25 28.8 10.8 1263 14 US-10-032-585-6711 Sequence 6711, Ap
C 26 28.8 10.8 1263 14 US-10-156-761-1370 Sequence 1370, Ap
C 27 28.8 10.8 28690 10 US-09-070-927A-138 Sequence 138, App
C 28 28.8 10.8 9025608 14 US-10-156-761-1 Sequence 1, Appl
C 29 28.4 10.6 476 11 US-09-918-395-20543 Sequence 20543, A
C 30 28.4 10.6 537 10 US-09-974-300-5520 Sequence 5520, Ap
C 31 28.4 10.6 1290 14 US-10-156-761-7250 Sequence 7250, Ap
C 32 28.4 10.6 27893 12 US-10-017-161-757 Sequence 757, App
C 33 28.2 10.6 269 10 US-09-878-574-8981 Sequence 8981, Ap
C 34 28.2 10.6 1188 11 US-09-752-957B-1 Sequence 1, Appl
C 35 28 10.5 231 9 US-09-864-761-27764 Sequence 27764, A
C 36 28 10.5 414 10 US-09-960-352-731 Sequence 731, App
C 37 28 10.5 469 9 US-09-864-761-11136 Sequence 11136, A
C 38 28 10.5 693 10 US-09-738-626-2956 Sequence 2956, Ap
C 39 28 10.5 2964 10 US-09-974-300-1533 Sequence 1533, Ap
C 40 28 10.5 8119 9 US-09-802-669-45 Sequence 45, Appl
C 41 28 10.5 8133 12 US-10-133-937-16 Sequence 16, Appl
C 42 27.8 10.4 524 12 US-10-027-632-324551 Sequence 324551, Sequence 324551,
C 43 27.8 10.4 524 13 US-10-027-632-324551 Sequence 208075,
C 44 27.8 10.4 546 12 US-10-027-632-208075 Sequence 208075,
C 45 27.8 10.4 546 13 US-10-027-632-208075 Sequence 208075,

ALIGNMENTS

RESULT 1

US-09-928-457-89

Sequence 89, Application US/09928457

Patent No. US20020164603A1

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: DNA, specific proteins and peptides

TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method

TITLE OF INVENTION: for obtaining them and their biological application.

NUMBER OF SEQUENCES: 99

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (OEB)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/928,457

FILING DATE: 2001-08-14

PRIOR APPLICATION DATA: 09/214,759

APPLICATION NUMBER: 09/214,759

FILING DATE: 199-12-10

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-928-457-89

Query Match 100.0%; Score 267; DB 10; Length 267;

Best Local Similarity 100.0%; Pred. No. 8.5e-84;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATGACACACGCATCATCGTTCCGGTGGCTTCGTTGGCATTAGCAGTTG 60

DB 1 AATTATGACACACGCATCATCGTTCCGGTGGCTTCGTTGGCATTAGCAGTTG 60

QY 61 CGGCTCAATCAATATGTAACGGTTTCGACAGAAATTCAGAAAGTGCCTGGTTTC 120
DB 61 CGGCTCAATCAATATGTAACGGTTTCGACAGAAATTCAGAAAGTGCCTGGTTTC 120
QY 121 CTTGGCGTCAACATGCGGTAAATAATCAGCAACGCGAGCAATGAAGCATACGCATCA 180
DB 121 CTTGGCGTCAACATGCGGTAAATAATCAGCAACGCGAGCAATGAAGCATACGCATCA 180
QY 181 CTTTACCGCAACTGTGGTGAAGCGGTGACCAATGCTATGTACCAAGTGAATCAGCACA 240
DB 181 CTTTACCGCAACTGTGGTGAAGCGGTGACCAATGCTATGTACCAAGTGAATCAGCACA 240
QY 241 ATCGCGGTACCACTTCGATGCAATT 267
DB 241 ATCGCGGTACCACTTCGATGCAATT 267

RESULT 2

US-10-044-090-506
; Sequence 506, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 506
; LENGTH: 1651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 403121.11
US-10-044-090-506

Query Match 13.0%; Score 34.6; DB 13; Length 1651;
Best Local Similarity 55.4%; Pred. No. 0.2;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGAACGTCGCGGTTTGCCTTGGCGTCAACAAATCCCGTAAAAATCAGCAACCGCAG 159
DB 207 TAAGGACCGCGCGGCTCTCACCGTGGCGTCAACCAACCGTGGCGAGTCCATCTCTCG 266
QY 160 CAATCAAGGCATACGCATCAACTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
DB 267 CGAGACCGACTGCGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACCA 326
QY 220 G 220
DB 327 G 327

RESULT 3

US-09-771-838A-17
; Sequence 17, Application US/09771838A
; Patent No. US20020028198A1
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,838A
FILING DATE: 29-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/182,983
FILING DATE: 1997-08-12
ATTORNEY/AGENT INFORMATION:
NAME: FETHERSTONHAUGH & Co.,
REFERENCE/DOCKET NUMBER: 28605-30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613)-232-2486
TELEFAX: (613)-232-8440
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-771-838A-17
Query Match 13.0%; Score 34.6; DB 9; Length 1796;
Best Local Similarity 55.4%; Pred. No. 0.21;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGAACGTCGCGGTTTGCCTTGGCGTCAACAAATCCCGTAAAAATCAGCAACCGCAG 159
DB 1319 TAAGGACCGCGCGGCTCTCACCGTGGCGTCAACCAACCGTGGCGAGTCCATCTCTCG 1378
QY 160 CAATCAAGGCATACGCATCAACTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
DB 1379 CGAGACCGACTGCGCGTCCACATCAACGAGCGCGGAGATCGCGTGGCCAGCACCA 1438
QY 220 G 220
DB 1439 G 1439
RESULT 4
US-09-771-838A-5
; Sequence 5, Application US/09771838A
; Patent No. US20020028198A1
; GENERAL INFORMATION:
; APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
; TITLE OF INVENTION: NOVEL GLUTAMINE:FRUCTOSE-6-PHOSPHATE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FETHERSTONHAUGH & Co.
; STREET: P.O. BOX 2999, STATION D
; CITY: OTTAWA
; STATE: ONTARIO
; COUNTRY: CANADA
; ZIP: K1P 5Y6
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,838A
FILING DATE: 29-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/182,983
FILING DATE: 1997-08-12
ATTORNEY/AGENT INFORMATION:

NAME: FETHERSTONHAUGH & Co.,
REFERENCE/DOCKET NUMBER: 28605-30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613)-232-2486
TELEFAX: (613)-232-8440
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-771-838A-5

Query Match 13.0%; Score 34.6; DB 9; Length 1845;
Best Local Similarity 55.4%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAAGTCGCGGTTTGCTTGGGGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 1311 TAAGGACCGCGGCTCTCACGGTGGGGGTCCACCAACCGTGGCGAGCTCCATCTCTCG 1370

QY 160 CAATGAAGCATACGATCACTTACCGCAACTGTGGTAAAGCGCGTGACCAATGCTAT 219
DB 1371 CGAGACCGACTCGCGGCTCCACATCAACGAGCGCGGAGATCGGCGTGCCAGCACCAA 1430

QY 220 G 220
DB 1431 G 1431

RESULT 5
US-09-771-838A-6
Sequence 6, Application US/09771838A
Patent No. US20020028198A1
GENERAL INFORMATION:
APPLICANT: TAKEDA CHEMICAL INDUSTRIES, LTD.
TITLE OF INVENTION: NOVEL GLUTAMINE-FRUCTOSE-6-PHOSPHATE
AMIDOTRANSFERASE, ITS PRODUCTION AND USE
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FETHERSTONHAUGH & Co.
STREET: P.O. BOX 2999, STATION D
CITY: OTTAWA
STATE: ONTARIO
COUNTRY: CANADA
ZIP: K1P 5Y6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,838A
FILING DATE: 29-Jan-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/182,983
FILING DATE: 1997-08-12
ATTORNEY/AGENT INFORMATION:
NAME: FETHERSTONHAUGH & Co.,
REFERENCE/DOCKET NUMBER: 28605-30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613)-232-2486
TELEFAX: (613)-232-8440
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2046 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-771-838A-6

Query Match 13.0%; Score 34.6; DB 9; Length 2046;
Best Local Similarity 55.4%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 100 TCAGGAAGTCGCGGTTTGCTTGGGGGTCCACCAATGCCGTAAATAATCAGCAACCGCAG 159
DB 1311 TAAGGACCGCGGCTCTCACGGTGGGGGTCCACCAACCGTGGCGAGCTCCATCTCTCG 1370

QY 160 CAATGAAGCATACGATCACTTACCGCAACTGTGGTAAAGCGCGTGACCAATGCTAT 219
DB 1371 CGAGACCGACTCGCGGCTCCACATCAACGAGCGCGGAGATCGGCGTGCCAGCACCAA 1430

QY 220 G 220
DB 1431 G 1431

RESULT 6
US-09-079-892-4
Sequence 4, Application US/09079892
Patent No. US20020061301A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Freeta
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN CARBOHYDRATE METABOLISM ENZYMES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,892
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0524 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3023 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SINTBST01
CLONE: 1429011
US-09-079-892-4

Query Match 13.0%; Score 34.6; DB 9; Length 3023;
Best Local Similarity 55.4%; Pred. No. 0.28;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGGACGTCGGCGTTGGCGTCAACCAATGCCGTAATAATCAGCAACCGCAG 159
DB 1409 TAAGGACCGCGCGCTCTCACCGTGGCGTCAACCAACCGCTGGCGAGCTCCATCTCTCG 1468
QY 160 CAATGAAGCATAGCGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
DB 1469 CGAGACCGACTGCGGCGTCCACATCAAGCGAGCGCGGAGATGGCGTGGCCAGCACAA 1528
QY 220 G 220
DB 1529 G 1529
RESULT 7
US-10-044-090-685
; Sequence 685, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 685
; LENGTH: 3048
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 403121.12
US-10-044-090-685

Query Match 13.0%; Score 34.6; DB 13; Length 3048;
Best Local Similarity 55.4%; Pred. No. 0.28; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 100 TCAGGACGTCGGCGTTGGCGTCAACCAATGCCGTAATAATCAGCAACCGCAG 159
DB 1429 TAAGGACCGCGCGCTCTCACCGTGGCGTCAACCAACCGCTGGCGAGCTCCATCTCTCG 1488
QY 160 CAATGAAGCATAGCGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCAATGCTAT 219
DB 1489 CGAGACCGACTGCGGCGTCCACATCAAGCGAGCGCGGAGATGGCGTGGCCAGCACAA 1548
QY 220 G 220
DB 1549 G 1549

RESULT 8
US-09-867-701-3070
; Sequence 3070, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3070
; LENGTH: 369
; TYPE: DNA

; ORGANISM: Homo sapien
US-09-867-701-3070
Query Match 11.8%; Score 31.4; DB 10; Length 369;
Best Local Similarity 54.9%; Pred. No. 1.3;
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 100 TCAGGACGTCGGCGTTGGCGTCAACCAATGCCGTAATAATCAGCAACCGCAG 159
DB 151 TAAGGACCGCGCGCTCTCACCGTGGCGTCAACCAACCGCTGGCGAGCTCCATCTCTCG 210
QY 160 CAATGAAGCATAGCGATCAACTTTACCGCAACTGTGGTAAGCGCGTGACCA 212
DB 211 CGAGACCGACTGCGGCGTCCACATCAAGCGAGCGCGGAGATGGCGTGGCCA 263
RESULT 9
US-10-027-632-59064
; Sequence 59064, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59064
; LENGTH: 1601042
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(1601042)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

Query Match 11.7%; Score 31.2; DB 12; Length 1601042;
Best Local Similarity 64.3%; Pred. No. 94;
Matches 45; Conservative 1; Mismatches 24; Indels 0; Gaps 0;
QY 126 GCGTCACCAATCCGTAATAATCAGCAACCGCAGCAATGAAGGCGATACGCACTTTA 185
DB 1005504 GCATTAATTAATCAATATATCAAGCAAGCAGCTATCACGTGATTGTTAACAATGA 1005563
QY 186 CCGCACTGT 195
DB 1005564 CCGCACTGT 1005573

RESULT 10
US-10-027-632-59064
; Sequence 59064, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59064
; LENGTH: 1601042
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(1601042)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59064
; LENGTH: 1601042
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1601042)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

Query Match      11.7%; Score 31.2; DB 13; Length 1601042;
Best Local Similarity 64.3%; Pred. No. 94;
Matches 45; Conservative 1; Mismatches 24; Indels 0; Gaps 0;

QY 126 GCGTCACCAATCCGCTAAATAATCAGCACCCAGCAATGAGGCATAGGCATCAACTTAA 185
DB 1005504 GCATAATTAATGCCAATAATATCAGCAAYAGCAGTATCAGTGATGTTTAAACAATAGA 1005553

QY 186 CGCAACTGT 195
DB 1005564 CGGCACGTGT 1005573

RESULT 11
US-10-037-270-914
; Sequence 914, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dairui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
```

```
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: PC_FL_genes Version 1.0
; SEQ ID NO 914
; LENGTH: 1466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (779)..(1213)
; NAME/KEY: misc feature
; LOCATION: (1)..(1466)
; OTHER INFORMATION: n = a,t,c or g
US-10-037-270-914

Query Match      11.4%; Score 30.4; DB 14; Length 1466;
Best Local Similarity 57.3%; Pred. No. 5.8;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 89 GACCAGAAACTTCAGAAAGCTGCGGGTTTGCTTTGGGCGTCACCAATCCCGTAAATAATC 148
DB 478 GAGGAGAAGCTTCGGAACACAGGCTGCTCCCAAGGGCGCTCAGTTTCCATTATATGTC 537

QY 149 AGCAACCGCAGCAATGAAGCATACCATCAACTTT 184
DB 538 AACACCAGCCTTGAAGACATGAGGAGCAACAACGTT 573

RESULT 12
US-10-175-523-67
; Sequence 67, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/1J795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 67
; LENGTH: 74868
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-67

Query Match      11.2%; Score 30; DB 14; Length 74868;
Best Local Similarity 61.5%; Pred. No. 56;
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 39 TTGGTTGGCATTAGCAGGTTGCGGTCATCAATCAATATGTAACCGTTTCGACCAAGAAC 98
DB 13595 TAGAGTTGGAGTGTGAAGATATAGCTTATTCAGATATCAAGCATCAAGCCATTGACGCGCAAA 13654

QY 99 TTCAGGAACGTGCCGGT 116
```

Db 13655 TTTAGAAAAGAGGCTGT 13672

RESULT 13

US-09-796-753-47

; Sequence 47, Application US/09796753

; Publication No. US20030027998A1

; GENERAL INFORMATION:

; APPLICANT: McCarthy, Sean A.

; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF

; FILE REFERENCE: 7853-227-999

; CURRENT FILING DATE: 2001-03-01

; CURRENT APPLICATION NUMBER: US/09/796,753

; PRIOR APPLICATION NUMBER: 09/183,175

; PRIOR FILING DATE: 1998-10-30

; PRIOR APPLICATION NUMBER: 09/223,094

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/223,546

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/224,246

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/259,388

; PRIOR FILING DATE: 1999-02-26

; PRIOR APPLICATION NUMBER: 60/122,458

; PRIOR FILING DATE: 1999-03-01

; PRIOR APPLICATION NUMBER: 09/312,359

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 09/336,536

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: 09/342,687

; PRIOR FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: 09/345,464

; PRIOR FILING DATE: 1999-06-30

; PRIOR APPLICATION NUMBER: 09/365,164

; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: 09/399,723

; PRIOR FILING DATE: 1999-09-20

; PRIOR APPLICATION NUMBER: 09/409,634

; PRIOR FILING DATE: 1999-09-30

; PRIOR APPLICATION NUMBER: 09/471,179

; PRIOR FILING DATE: 1999-12-23

; PRIOR APPLICATION NUMBER: 09/474,071

; PRIOR FILING DATE: 1999-12-29

; PRIOR APPLICATION NUMBER: 09/474,072

; PRIOR FILING DATE: 1999-12-29

; PRIOR APPLICATION NUMBER: 09/514,010

; PRIOR FILING DATE: 2000-02-25

; PRIOR APPLICATION NUMBER: 09/516,745

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 09/572,002

; PRIOR FILING DATE: 2000-05-14

; PRIOR APPLICATION NUMBER: 09/597,993

; PRIOR FILING DATE: 2000-06-19

; PRIOR APPLICATION NUMBER: 09/599,596

; PRIOR FILING DATE: 2000-06-22

; PRIOR APPLICATION NUMBER: 09/630,334

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: 09/606,565

; PRIOR FILING DATE: 2000-06-29

; PRIOR APPLICATION NUMBER: 09/606,317

; PRIOR FILING DATE: 2000-06-29

; PRIOR APPLICATION NUMBER: 09/665,666

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: 09/677,751

; PRIOR FILING DATE: 2000-09-30

; NUMBER OF SEQ ID NOS: 162

; SEQ ID NO 47

; LENGTH: 1473

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-796-753-47

Query Match 11.0%; Score 29.4; DB 11; Length 1473;

Best Local Similarity 56.8%; Pred. No. 13;

Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 18 TCATCGTTTCGGCTCGGTTGGTTGGCATTAGCAGGTTGGCGCTCAATCAATAATG 77

Db 93 TCGCAGTTTCTGCATTCTGTGGGGCGGTGAGCAATCGACGGCTGGCTCAGTCAATAAG 152

QY 78 TAACCGTTTCCGACGAGAAACTTCAGGAACGTGCC 112

Db 153 GATCCGGGCCCAACACGACATTCACAGTAGTTC 187

RESULT 14

US-10-198-846-11017

; Sequence 11017, Application US/10198846

; Publication No. US20030099974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11017

; LENGTH: 1481

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1469, 1470, 1471, 1472, 1473, 1474, 1475, 1476, 1477, 1478,

; LOCATION: 1479, 1480, 1481

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-11017

Query Match 11.0%; Score 29.4; DB 14; Length 1481;

Best Local Similarity 56.8%; Pred. No. 13;

Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 18 TCATCGTTTCGGCTCGGTTGGTTGGCATTAGCAGGTTGGCGCTCAATCAATAATG 77

Db 100 TCGCAGTTTCTGCATTCTGTGGGGCGGTGAGCAATCGACGGCTGGCTCAGTCAATAAG 159

QY 78 TAACCGTTTCCGACGAGAAACTTCAGGAACGTGCC 112

Db 160 GATCCGGGCCCAACACGACATTCACAGTAGTTC 194

RESULT 15

US-10-098-841-67

; Sequence 67, Application US/10098841

; Publication No. US20020197679A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Xu, Chongjun

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunging

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhao, Qing A.

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: Wehrman, Tom

Search completed: November 15, 2003, 08:32:28
Job time : 163.297 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 1145.51 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-92
Perfect score: 259
Sequence: 1 AATTGCTGATATCAAGTAGG.....TACTTACTGCGAGCGAAATT 259

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_in.*
18: em_mu.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	100.0	259	1	AF169468 Neisseria
2	259	100.0	259	6	A68921 Sequence 92
3	259	100.0	259	6	BD03021 DNA and s
C 4	255.8	98.8	349061	1	AL162753 Neisseria
5	56.8	21.9	1275	6	A96274 Sequence 30
6	52.6	20.3	1572	6	AX024062 Sequence
7	52.6	20.3	1575	6	A96276 Sequence 30
8	52.6	20.3	1575	6	AX024113 Sequence
9	50.4	19.5	15053	1	AE002550 Neisseria
10	50.4	19.5	172325	6	AX044035 Sequence
11	50.4	19.5	349980	6	AX044034 Sequence
C 12	41.2	15.9	13326	6	AX346615 Sequence
C 13	41	15.8	90284	8	T2K10
14	40.6	15.7	236159	2	AC125553
15	40.6	15.7	253253	2	AC126527
16	40	15.4	222039	2	AC107506
C 17	39.2	15.1	173138	5	AC108132
C 18	39.2	15.1	180964	2	AC025846
19	39.2	15.1	217384	2	AC023573
C 20	39	15.1	327738	2	AC104922
21	38.6	14.9	221866	2	AC114426
22	38.4	14.8	91940	9	AL445674
23	38.2	14.7	15282	6	AX251836
C 24	38.2	14.7	15282	6	AX348621
C 25	38	14.7	6944	6	AX347467
C 26	38	14.7	6944	6	AX349188
C 27	38	14.7	6944	6	AX657867
C 28	38	14.7	6944	6	AX659141
C 29	38	14.7	16509	6	AX346222
C 30	37.8	14.6	5552	6	AX346161
31	37.6	14.5	1576	6	A96272
32	37.6	14.5	1576	6	A96270
33	37.2	14.4	94056	9	AL139192
34	37.2	14.4	146753	9	AC105753
C 35	37.2	14.4	153615	2	AC111043
36	37.2	14.4	181443	9	AC099775
37	37	14.3	3399	6	BD113032
C 38	37	14.3	95976	9	AC068062
C 39	37	14.3	349980	6	AX344560
40	36.8	14.2	35301	8	AP002544
41	36.8	14.2	219340	10	AL844881
C 42	36.6	14.1	8576	6	AX347130
C 43	36.4	14.1	2200	6	AX347092
C 44	36.4	14.1	6239	6	AX251987
C 45	36.4	14.1	6239	6	AX344383

ALIGNMENTS

RESULT 1
AF169468 AF169468 259 bp DNA linear BCT 09-AUG-2000
LOCUS Neisseria meningitidis strain Z2491 clone Em045 unknown sequence.
DEFINITION
ACCESSION AF169468
VERSION AF169468.1 GI:9754676
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 259)
Perrin A., Nassif X. and Tinsley C.R.
AUTHORS Identification of regions of the chromosome of Neisseria
TITLE meningitidis and Neisseria gonorrhoeae which are specific to

Pred. No. is the number of results predicted by chance to have a

```

pathogenic Neisseriae
Unpublished
REFERENCE 2 (bases 1 to 259)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        1..259
        /organism="Neisseria meningitidis"
        /mol_type="genomic DNA"
        /strain="22491"
        /db_xref="taxon:487"
        /clone="Em045"
BASE COUNT 81 a 71 c 34 g 73 t
ORIGIN
Query Match 100.0%; Score 259; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 4.7e-65;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
DB 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
QY 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
DB 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
QY 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
DB 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
QY 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
DB 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
QY 241 ACTTACTGCCAGCAAAATT 259
DB 241 ACTTACTGCCAGCAAAATT 259

RESULT 3
LOCUS BD063021
DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications.
ACCESSION BD063021
VERSION BD063021.1 GI:22608624
KEYWORDS JP 2001504684-A/83.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 259)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 83 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/83
PD 10-APR-2001
PF 11-JUL-1997 JP 19980505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..259
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 81 a 71 c 34 g 73 t
ORIGIN
Query Match 100.0%; Score 259; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 4.7e-65;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
DB 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
QY 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
DB 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
QY 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
DB 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
QY 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
DB 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
QY 241 ACTTACTGCCAGCAAAATT 259
DB 241 ACTTACTGCCAGCAAAATT 259

RESULT 2
LOCUS A68921
DEFINITION Sequence 92 from Patent WO9802547.
ACCESSION A68921
VERSION A68921.1 GI:4759840
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 259)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
Merker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
APPLICATIONS
JOURNAL Patent: WO 9802547-A 92 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
COMMENT Other publication FR 2751000 19980116.
FEATURES
    source
        1..259
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 81 a 71 c 34 g 73 t
ORIGIN
Query Match 100.0%; Score 259; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 4.7e-65;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
DB 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
QY 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
DB 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
QY 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
DB 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
QY 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
DB 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
QY 241 ACTTACTGCCAGCAAAATT 259
DB 241 ACTTACTGCCAGCAAAATT 259

RESULT 3
LOCUS BD063021
DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications.
ACCESSION BD063021
VERSION BD063021.1 GI:22608624
KEYWORDS JP 2001504684-A/83.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 259)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and
Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
species bacteria, method for obtaining them and their biological
applications
JOURNAL Patent: JP 2001504684-A 83 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/83
PD 10-APR-2001
PF 11-JUL-1997 JP 19980505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31,C07K14/22,C07K16/12,A61K39/095,C12Q1/68,G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..259
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 81 a 71 c 34 g 73 t
ORIGIN
Query Match 100.0%; Score 259; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 4.7e-65;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
DB 1 AATTGTGTATATCAAGTAGGATGGCGCATTTATGCTGACCTACAAACCAAAACCAACCT 60
QY 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
DB 61 ACCACCCCTTAATCAACTCCACAAACCCCTCTTCAGACAAACCTCGTTTTTGGAAAAACAATC 120
QY 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
DB 121 TGTAAACAGATAACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTT 180
QY 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
DB 181 TTATTGTGAACCTTCCCATTTATGAGAAAATCCCTTTTCGTCTCTCTTTCTGTATTGTCCTCCT 240
QY 241 ACTTACTGCCAGCAAAATT 259
DB 241 ACTTACTGCCAGCAAAATT 259

```

```

QY 121 TGTAAACAGATAACTGCTGAGATACCGTTGCGAGCCGACCAAAACCCGTACTGCAACTT 180
Db 121 TGTAAACAGATAACTGCTGAGATACCGTTGCGAGCCGACCAAAACCCGTACTGCAACTT 180
QY 181 TTATTGTGAACCTCCCAATTATGAGAAAATCCCTTTTCGTCCTCTCTTTCTGTATTGTCGCCCT 240
Db 181 TTATTGTGAACCTCCCAATTATGAGAAAATCCCTTTTCGTCCTCTCTTTCTGTATTGTCGCCCT 240
QY 241 ACTTACTGCCAGCGAAATT 259
Db 241 ACTTACTGCCAGCGAAATT 259

RESULT 4
NMA222491/c NMA222491 349061 bp DNA linear BCT 02-SEP-2002
LOCUS Neisseria meningitidis serogroup A strain Z2491 complete genome;
DEFINITION segment 2/7.
ACCESSION AL162753 AL157959
VERSION AL162753.2 GI:7379120
SOURCE Neisseria meningitidis Z2491
ORGANISM Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 349061)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
Holroyd,S., Jagsis,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
PubMed 10761919
REFERENCE 2 (bases 1 to 349061)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
NOTES Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source 1..349061
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/notes="serogroup: A"
complement(24..206)
/notes="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
310.39"
/label=ATR
RBS 209..212
Gene 220..681
CDS 220..681
/gene="NMA0368"
/gene="NMA0368"
/notes="NMA0368", probable integral membrane protein, len:
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83669.1"
/db_xref="GI:7379121"
/db_xref="SPTREMBL:Q9JWH7"
/translation="MQQNRKSPFIVMLLVSVLWIAISLVNVAFYIGNHGSMEGLTV

```

gene

CDS

```

LILGSIPLASDIRCAVYVWLAAILVLLALRKVVPVHFAFWCLALVALVSKAVVY
DEAGNTSDIVRYGAGFYLYWAFASVAGTGTAGKXKKAASADGTDNDV"
671..1492
/gene="hemk"
671..1492
/gene="hemk"
/notes="NMA0369, hemk, HemK protein, len: 273 aa; similar
to e.g. HEMK ECOLI P37186 HEMK protein (277 aa), fasta
scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
/codon_start=1
/transl_table=11
/product="HemK protein"
/protein_id="CAB83670.1"
/db_xref="GI:7379122"
/db_xref="SPTREMBL:Q9JWH6"
/translation="MFDKWLGLSKLPKXNEARMMLQYVSEYTRVOLLTRGGEEMPDEV
QRADLRQRLNGEPVAYILGAREFFGRFTVNPVSLIPRTEHLVEAVLARLPEN
GRWDLGTGSAVAVTVALERPDFAFRASDI SPPALETARKAADLGRVAFYGSWF
DTDMPSECKWDIIIVSNPPIYINGDKHLQSGDLRFEPQIALTDFDGLSCIRTLAQGAP
DRLAEGGFLLEHGFDQGAARGVLAENGPSGVETLPDLAGLDRLVTLCGYMKHLK"
831..840
/gene="hemk"
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
1162..1171
/gene="hemk"
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
1187..1207
/gene="hemk"
/notes="hemk"
/notes="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
1565..1568
1578..1596
/gene="NMA0370"
1578..1596
/gene="NMA0370"
/notes="NMA0370"
/notes="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
P44640 hypothetical protein H30325 (450 aa), fasta scores;
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83671.1"
/db_xref="GI:7379123"
/db_xref="SPTREMBL:Q9JWH5"
/translation="MNAVVAIVMLVLSLRVHVVLSTIGAFVGGAVAGMPLQNI
DAAGVQAGIIPVFNKLEGAKIALSYAMLGAFAMATHSGLPQOLAGAVVRKLN
GMPDSVRSRGVVKWLLSILVNMMSQNIPIPIAFIPVLPPLLVFNRLKIDR
RLIACVITFGLVTVFPLPGFALFLNEILLGNHSAAPOLDVKNVNMAMAPAL
CMLAGLLAFVHYKRPRLYQSNADTACNADANRPOPSVRSIAAVALVAFPAIQL
MYEDSELVGMGLGFVFMGLVGNNDKANDVFGEGIKMMVGFIMIAQSPFAVMA
TGHIOPLVSSNAIFGNSKMAALVVLVGLVTMGISFSSTLPIIAIYVPLCVGL
GFSPLATVAIVGTAGALDAGSPASDSTLGTMLGNADQGDHIDRHSVPTFIHYNIP
2967..3191
/gene="slyX"
2967..3191
/gene="slyX"
/notes="NMA0371, slyX, slyX protein homolog, len: 74 aa;
similar to SLYX_HAEIN P44759 SLYX protein homolog (73 aa),
fasta scores; E(): 0.0047, 33.8% identity in 74 aa
overlap, and SLYX_ECOLI P30857 SLYX protein (72 aa), fasta
scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=11
/product="slyX protein homolog"
/protein_id="CAB83672.1"

```

gene

CDS

```

/db_xref="GI:7379124"
/db_xref="SPTRMBL:O9JWH4"
/translation="MDAVQSFERRHTELETSQALQEDVIAGLNAMVAELRQTLDLQQA
QRLLYQMDRNPDAQEPYSLRDEIPPH"
complement(3271..3423)
/gene="NMA0372"
CDS
complement(3271..3423)
/gene="NMA0372"
note="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=1
/product="very hypothetical protein NMA0372"
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="SPTRMBL:O9JWH3"
/translation="MGRSLPLSGDGFLLAVYEPVGFICLQKGTAYFRSGGKRFHRI
KGLUSD"
3562..3571
misc_feature
/label="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3586..4356)
/gene="thiF"
CDS
complement(3586..4356)
/gene="thiF"
note="NMA0373, thiF, probable ThiF protein, len: 256 aa;
similar to e.g. THIF_ECOLI P30138 ThiF protein (251 aa),
fasta scores; E(): 0.43.1% identity in 246 aa overlap,
and MOBE_ECOLI P12282 molybdopter in biosynthesis MOBE
protein_(249 aa), fasta scores; E(): 0.43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopter in biosynthesis proteins). Contains
pfam match to entry PF00899 ThiF_family, ThiF family"
/codon_start=1
/transl_table=1
/product="ThiF protein"
/protein_id="CAB83674.1"
/db_xref="GI:7379126"
/db_xref="SPTRMBL:O9JWH2"
/translation="MTTEHDNDLFLRYSRHILDEIGEQQLSAAHLVVGCG
GLGAALPYLAAGIGTLTADSDTVLHNLQRFVDEGDKLTELADLRHIN
HTVDVRIKGLDGLTGVQADIVLDCDNYATQVNRACVQAKTPLYSGAAVR
FEGSLAVRDLDPSPCYACLDGSGASDGLSGVSPVGLVIGISTQABALKILL
DAGEPSGRILAVRALGEGQYFDLPRNEPCVCGAER"
3787..3796
misc_feature
/label="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3847..4257)
/gene="thiP"
note="Pfam match to entry PF00899 ThiF_family, ThiF
family, score 186.60, E-value 4.1e-52"
complement(4418..4427)
/misc_feature
/label="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4471..7224
/gene="ppc"
4471..7224
/gene="ppc"
/EC_number="4.1.1.31"
note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
917 aa; similar to many e.g. CAPP_RHOPA O32483
phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
fasta scores; E(): 0.43.3% identity in 928 aa overlap.
Contains 2x pfam match to entry PF00311 PEPCase,
phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
site motif A (P-loop), and PS00393 Phosphoenolpyruvate
carboxylase active site 2"
/codon_start=1
/transl_table=1
/product="phosphoenolpyruvate carboxylase"
/protein_id="CAB83675.1"
Query Match 98.8%; Score 255.8; DB 1; Length 349061;
Best Local Similarity 99.2%; Pred. No. 3.9e-64;

```

```

Matches 257; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATTGTGTATATCAAGTAGGATGGCATTATGCTGACCTACAAAACCAAAACAACTT 60
|||||
Db 65981 AATTGTGTATATCAAGTAGGATGGCATTATGCTGACCTACAAAACCAAAACAACTT 65922
|||||
QY 61 ACCACCCCTTAATCAACTCCACAAAACCTCTTCAGACAACCTCGTTTTTGA AAAACAATC 120
|||||
Db 65921 ACCACCCCTTAATCAACTCCACAAAACCTCTTCAGACAACCTCGTTTTTGA AAAACAATC 65862
|||||
QY 121 TGTAAACAGATAACTGTGGAAGATACCGTTGCCGAGCCCAAAACCCGTAACATTT 180
|||||
Db 65861 TGTAAACAGATAACTGTGGAAGATACCGTTGCCGAGCCCAAAACCCGTAACATTT 65802
|||||
QY 181 TTATTGTCAACTTCCATTATGAGAAAATCCCTTTTCGTCCTCTTTCTGTATTCTGTCCT 240
|||||
Db 65801 TTATTGTCAACTTCCATTATGAGAAAATCCCTTTTCGTCCTCTTTCTGTATTCTGTCCT 65742
|||||
QY 241 ACTTACTCCAGCGAAATTT 259
|||||
Db 65741 ACTTACTCCAGCGAAATTT 65723
|||||
RESULT 5
A96274 1275 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 307 from Patent WO9924578.
DEFINITION A96274
ACCESSION A96274
VERSION A96274.1 GI:5780029
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Pizza M., Scarlato V., Rappuoli R., Grandi G. and Massignani V.
TITLE Neisserial antigens
JOURNAL Patient: WO 9924578-A 307 20-MAY-1999;
PIZZA MARIA GRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); Massignani VEGA (IT)
FEATURES
Location/Qualifiers
source
1..1275
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 300 a 304 c 319 g 352 t
ORIGIN
Query Match 21.9%; Score 56.8; DB 6; Length 1275;
Best Local Similarity 96.7%; Pred. No. 8.5e-06;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 200 ATGAGAAAATCCCTTTTCGTCCTCTTTCTGTATTCTGTCCTTACTTACTGCGAGGAAATTT 259
|||||
Db 1 ATGAAAAATCCCTTTTCGTCCTCTTTCTGTATTCTGTCCTTACTTACTGCGAGGAAATTT 60
|||||
RESULT 6
AX024062 1572 bp DNA linear PAT 15-SEP-2000
LOCUS Sequence 5 from Patent FR2785293.
DEFINITION AX024062
ACCESSION AX024062
VERSION AX024062.1 GI:10184374
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
REFERENCE 1
AUTHORS Nassif X., Tinsley C., Autame L., Perrin A., Rokbi B.,
Bouchardon A. and Renaud M.G.
JOURNAL Patent: FR 2785293-A 5 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES
Location/Qualifiers

```

```

source
1..1572
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
1..1572
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC08907.1"
/db_xref="GI:10184375"
/translation="MKKSLFVFLYSSLLTASEIAYRVFGIETLPAKVAETFAITF
MIAALYLPARYKASRLIIAVFAFSAIANNVHYAVYQSMGTGINYKLMKEITRVGSA
GASLDKLMWLPALWGVAEVMFLFCSLAKFRKTHFSADILPAFLMLMFVRSFDTKQEH
GISEPKTYSRIKANYFSFGYFVGRVLPYQLFDLSKIPVFKQAPSKIGQSGVQNVILI
MGESAAHLKLFYGRGTSPTLRLSQADFPKIVKQSYAGFMATVSLFSPFNVLPH
ANGLEQISGGDTNMFRLAKEQGYETFYSAQAENEMAILNLIGKWDHILQIOTQLGY
GNGDNMPDEKLLPLPDKINLQGRHFIVLHQRGHAPYVALLQPDQVFGELIVDKDY
DNTIHKTDOMIOTVFELQKQPDGNWLFAVTSDHQYVRQDIYNQGTVPQPSYLVPLV
YSSNKAVQQAANQAPFAPEIAFHQQLSTFLHTLGYDMPVSGCREGSGVTGNLTGDAG
SLNTRDGAKEYVYPO"
BASE COUNT      380 a   383 c   395 g   414 t
ORIGIN
Query Match      20.3%; Score 52.6; DB 6; Length 1572;
Best Local Similarity 93.2%; Pred. No. 0.00014;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 200 ATGAGAAATCCCTTTTCGTCCTCTTCTGTATTCGTCCTACTTACTGCGAGCGAAAT 258
|||||
Db 1 ATGAAAAATCCCTTTTCGTCCTCTTCTGTATTCATCCTACTTACCGCGAGCGAAAT 59
|||||

RESULT 7
A96276 A96276 1575 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 309 from Patent WO924578.
DEFINITION Neisserial antigens
ACCESSION Patent: WO 924578-A 309 20-MAY-1999.
VERSION PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
KEYWORDS CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
SOURCE Location/Qualifiers
ORGANISM 1..1575
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT      385 a   387 c   392 g   411 t
ORIGIN
Query Match      20.3%; Score 52.6; DB 6; Length 1575;
Best Local Similarity 93.2%; Pred. No. 0.00014;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 200 ATGAGAAATCCCTTTTCGTCCTCTTCTGTATTCGTCCTACTTACTGCGAGCGAAAT 258
|||||
Db 1 ATGAAAAATCCCTTTTCGTCCTCTTCTGTATTCATCCTACTTACCGCGAGCGAAAT 59
|||||

RESULT 8
AX024113 AX024113 1575 bp DNA linear PAT 15-SEP-2000
LOCUS Sequence 56 from Patent FR2785293.
DEFINITION Neisserial antigens
ACCESSION AX024113
VERSION AX024113.1 GI:10184424
KEYWORDS Neisseria gonorrhoeae
SOURCE Pizza M., Scarlato, V., Rappuoli, R., Grandi, G. and Massignani, V.
Neisserial antigens
Patent: WO 924578-A 309 20-MAY-1999.
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
LOCATION/QUALIFIERS
1..1575
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT      385 a   387 c   392 g   411 t
ORIGIN
Query Match      20.3%; Score 52.6; DB 6; Length 1575;
Best Local Similarity 93.2%; Pred. No. 0.00014;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 200 ATGAGAAATCCCTTTTCGTCCTCTTCTGTATTCGTCCTACTTACTGCGAGCGAAAT 258
|||||
Db 1 ATGAAAAATCCCTTTTCGTCCTCTTCTGTATTCATCCTACTTACCGCGAGCGAAAT 59
|||||

ORGANISM
Neisseria gonorrhoeae
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
Bouchardon, A. and Renaud, M.G.
Patent: FR 2785293-A 56 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)
JOURNAL
FEATURES
Location/Qualifiers
1..1575
/organism="Neisseria gonorrhoeae"
/mol_type="genomic DNA"
/db_xref="taxon:485"
1..1575
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC08932.1"
/db_xref="GI:10184425"
/translation="MKKSLFVFLYSSLLTASEIAYRVFGIETLPAKVAETFAITF
MIAALYLPARYKASRLIIAVFAFSAIANNVHYAVYQSMGTGINYKLMKEITRVGSA
GASLDKLMWLPALWGVAEVMFLFCSLAKFRKTHFSADILPAFLMLMFVRSFDTKQEH
GISEPKTYSRIKANYFSFGYFVGRVLPYQLFDLSKIPVFKQAPSKIGQSGVQNVILI
MGESAAHLKLFYGRGTSPTLRLSQADFPKIVKQSYAGFMATVSLFSPFNVLPH
ANGLEQISGGDTNMFRLAKEQGYETFYSAQAENEMAILNLIGKWDHILQIOTQLGY
GNGDNMPDEKLLPLPDKINLQGRHFIVLHQRGHAPYVALLQPDQVFGELIVDKDY
DNTIHKTDOMIOTVFELQKQPDGNWLFAVTSDHQYVRQDIYNQGTVPQPSYLVPLV
YSSNKAVQQAANQAPFAPEIAFHQQLSTFLHTLGYDMPVSGCREGSGVTGNLTGDAG
SLNTRDGAKEYVYPO"
BASE COUNT      385 a   387 c   392 g   411 t
ORIGIN
Query Match      20.3%; Score 52.6; DB 6; Length 1575;
Best Local Similarity 93.2%; Pred. No. 0.00014;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 200 ATGAGAAATCCCTTTTCGTCCTCTTCTGTATTCGTCCTACTTACTGCGAGCGAAAT 258
|||||
Db 1 ATGAAAAATCCCTTTTCGTCCTCTTCTGTATTCATCCTACTTACCGCGAGCGAAAT 59
|||||

RESULT 9
AE002550 AE002550 15053 bp DNA linear BCT 25-MAY-2000
LOCUS Neisseria meningitidis serogroup B strain MC58 section 192 of 206
DEFINITION of the complete genome.
ACCESSION AE002550 AE002098
VERSION AE002550.2 GI:7413476
KEYWORDS Neisseria meningitidis MC58
SOURCE Neisseria meningitidis MC58
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 15053)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, B.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Citron, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Massignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
JOURNAL MEDLINE
20175755
10710307
REFERENCE PUBMED
2 (bases 1 to 15053)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, B.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,

```

Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
 Cifton, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
 Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
 Pizzia, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
 Rappuoli, R. and Venter, J.C.
 Direct Submission
 Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 On Apr 4, 2000 this sequence version replaced gi:7227263.
 Location/Qualifiers
 source
 1..15053
 /organism="Neisseria meningitidis MC58"
 /mol_type="genomic DNA"
 /strain="MC58"
 /db_xref="taxon:122586"
 /note="serogroup: B"
 complement(74..982)
 /gene="NMB2000"
 complement(74..982)
 /gene="NMB2000"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42327.1"
 /db_xref="GI:7227264"
 /translation="MNQAINRADVRTRFIDDMVPLVGLHVLNVMQHVQKQVPA
 AIRRALGSLAAGLVLSGNLKNRGLTIVQGRGRKLMLVAASDRVTRATRWDET
 HEVLDLGLGEGVFLTLQKDEGQWGVPLEGGIAQMLVMKRSQELDT
 AIVLSADEAGGLVORLPEVDEAWHVSTLARTLAEELAGLDAQVLYRLPH
 BTPRVPEPEFESCRCRSKGVSDMLLGLGEEVGVVQGSIEVDCCFCHSKYVF
 DETDYNALPGEDVGVAKLPHRVQ"
 1227..1982
 /gene="NMB2001"
 1227..1982
 /note="NMB2001"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42328.1"
 /db_xref="GI:7227265"
 /translation="MDSFFKPAVWVLMFAVRPALADELTNLLSREOILRQFAED
 EQVLPINRPARAGNADELIGSAMGLNEQVLPVNRVPRBAGNADELIGNAMGLN
 EQVLPINRPARAGNADELIGSAMGLGIAIRYGTSTVSTGDCSGFQWHLFKRAM
 GINLPRTAEQRMGTIPVAKSELQFGDMVFFRILGGSKRISHVGLYIGNNRFTHAPRTG
 KNEITSLSHKYSKGVAFARVRKNDPSRFLN"
 2001..2219
 /gene="NMB2002"
 2001..2219
 /note="NMB2002"
 /note="hypothetical protein; identified by Glimmer2;
 putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAF42329.1"
 /db_xref="GI:7227266"
 /translation="MSNPENPKWYDDGQIVSCTEKVKVMSENVAELYQTAQAFEDA
 LLMCGGRQNDYLLALIEGLENPYRKV"
 2364..2711
 /gene="NMB2003"
 2364..2711
 /note="NMB2003"
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42330.1"
 /db_xref="GI:7227267"

/translation="NNIIRALLIILGCLATGTAVFLAGIKLPGSIVGMGLFALLQA
 GVKVTSWQLQTLDAALMSNLTLFLVPCVAVISYLDLIADDFWFSILVASASTLCVLLV
 TKVHRWRGIIR"
 2708..3400
 /gene="NMB2004"
 2708..3400
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF42331.1"
 /db_xref="GI:7227268"
 /translation="MNEILRQPSVLLFLTLAVYALAIIVTRTGNIFCPNVLSTIVL
 TAYKILGIDYAVYHNAQCFIDFWLKPAAVVLAVPLVQNRKIFNQMLPVLVSQIAGS
 VTGIVTGMVFAKWLGAEREVLSLASKSVNPIAIEITRSIGGIPATIAATVITAGLV
 GOIAGYKMLKNTVMPESSVGNLSLGTASHAMGIAASLERSRMAAAGLGLTFNGVLTA
 LIAPLLIPVLGF"
 3467..4687
 /gene="NMB2005"
 3467..4687
 /note="NMB2005"
 /note="similar to GB:M65216 SP:P38434 GB:S85363 PID:150243
 PID:246444 percent identity: 98.77; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="glutamate N-acetyltransferase; amino-acid
 acetyltransferase"
 /protein_id="AAF42332.1"
 /db_xref="GI:7227269"
 /translation="MAVNLTEKTAEPDIDGIALYTAQAGVKXPGHTDILLIAVAG
 STVGAVFTNRPAAVHAKSHUFDEGVFALVINTGNAGTCAQGRIDALAVCAA
 AARQIGCKPNQVLPSTGVIIEPLPMATIIAALPKMQPAPFNAARIMTDTVPKAA
 SREGKVGDKHTVATGIAGKSGMHPNMATLGTATDAKVSQVQLMTQEIADETF
 NITITVDGDTSTNDSFVIATGKNSQSEIDNIDPRYAQLKELLSLALEQAIVRQD
 EGATFTIVRVENAKTRDEARQAAVAVARSPLVKTFAPASDPNGLRLAAIGYAGVAD
 LDTDLAVNYLDDILLVAEHGGRASYTEAQGAVMSKAEITVRIKLHRGQAATVYTC
 LSHGVYSINADYRS"
 4747..6156
 /gene="NMB2006"
 4747..6156
 /note="NMB2006"
 /note="similar to PID:1652876 percent identity: 52.04;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="chloride channel protein-related protein"
 /protein_id="AAF42333.1"
 /db_xref="GI:7227270"
 /translation="MTSTFPRRLARKIOTRRLSRKSIAPLFLLAGSALVALTALFFA
 HLAQPALENAKLQVQVFWFAVVALPLGLPIIAMLTREKAPFTAGSGIPQVIAISLPL
 YDAQTRILRIGQLLKIPLTFGLMFLFCASIGREGSPSVQVGAAYVAGWAGACKXKGLA
 PKGQENDLMAAGAGLAAAFNAPLAGVIFAIELGRGIMLRWERQILLUGVLASGFI
 QVAIQGNPNFYSGNGVGLHEIFLWVALSGVGAAGLFGRLIYRGAAPFAPKIRG
 FIRNRPILLALMLGLLALLGTGYGTGTGYGHEAAQALHYEAPFGLAACKWLAT
 VFSYVAGVPGGIFTPTSLITGAVLGHETAAIDISQGANIIVLICMAAPLAGATQSPIT
 SAVVMTMTGGSLFLFWMLIACIFASQVSRQFSRPFVHASMRFQRQLQETAQTG
 NAPAPQATNSKTCMPSEN"
 6346..7737
 /gene="NMB2007"
 /note="This region contains a match to at least one other
 gene that is not full length, and is not the result of a
 sequencing artifact; similar to GP:1742299 percent
 identity: 79.36; identified by sequence similarity;
 putative"
 7881..9524
 /gene="NMB2008"
 7881..9524
 /note="similar to PID:862629 percent identity: 53.03;
 identified by sequence similarity; putative"

```

/codon_start=1
/transl_table=11
/product="ABC transporter, ATP-binding protein-related
protein"
/protein_id="AAF42334.1"
/db_xref="GI:727271"
/translation="MDEIQIPKVELOTKLENIKVLKSGSTIIIVANGTGKRLAV
YIEOLKEKAHRIQAHKRALKNPNVKIPEKSAKTYLSYGMWDGIQVSNRKNYRWN
NSYTHLNDPFOILLOYLPAQNNITAVANNOKLNENKVTNSKTKLDILOQAEWELLPH
RKLHITADIOVADVNDRELISANMSDGERALFYLGQVLSVDGSLVLPDEPELHI
KLSIISNWDKIEELRPPCSFLIITHDIEFAATVAKVYIRNYPTPAMDISEVPEP
NFDEETIMILGSRKPLFVFNENNSDIATRYCYPDWTIIPKACKDVIQSVSSILK
KLSNEMPLINKCSGIVLDSRDERIEQLNNLGIYILPVEIENLFLSLTDVAKEILK
LNQVSDLELLAKNEISLTLEQKJETWSEIKNEIQRCIQOQDLKLLTIYDNKGLLAKA
CVLGMKRNKHFESWMLTKGRNKFIDAIRQLPILD"
9568..9750
/gene="NMB2009"
/organism="Neisseria meningitidis"
/note="This region contains a gene with one or more
premature stops or frameshifts, and is not the result of a
sequencing artifact; similar to SP:Q09530 PID:733602
percent identity: 61.70; identified by sequence
similarity; putative"
9962..11539
gene

Query Match 19.5%; Score 50.4; DB 1; Length 15053;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 176 AACTTTTATGTGAACCTCCCATATGAGAAAATCCCTTTTCGTCCTTTCTGTATTCG 235
DB 9941 AATTTTATGTGAATAATTAATGATGAAAAAATCTTTCCTACGCTTGTTCTGTATTCG 10000
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 10001 TCCTTACTTACCGCAGCGAAATT 10024
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 10001 TCCTTACTTACCGCAGCGAAATT 10024

RESULT 10
LOCUS AX044035 172325 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 114 from Patent WO0066791.
ACCESSION AX044035
VERSION AX044035.1 GI:11342919
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1
AUTHORS Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 114 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Location/Qualifiers
source
1..172325
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 bases-seq 108: 300001 to 649980 bases-seq 110:
900001 to 1249980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 bases-seq
113: 1800001 to 2149980 bases-seq 114: 2100001 to
2272325 bases"
BASE COUNT 43072 a 47583 c 41465 g 40205 t
ORIGIN
Query Match 19.5%; Score 50.4; DB 6; Length 172325;
Best Local Similarity 75.0%; Pred. No. 0.00062;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 176 AACTTTTATGTGAACCTCCCATATGAGAAAATCCCTTTTCGTCCTTTCTGTATTCG 235
DB 323053 AATTTTATGTGAATAATTAATGATGAAAAAATCTTTCCTACGCTTGTTCTGTATTCG 323112
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 323113 TCCTTACTTACCGCAGCGAAATT 323136
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 323113 TCCTTACTTACCGCAGCGAAATT 323136

RESULT 12
LOCUS AX346615/c 13326 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1686 from Patent WO0200928.
ACCESSION AX346615
VERSION AX346615.1 GI:18494501
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1686 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers

```

```

Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 176 AACTTTTATGTGAACCTCCCATATGAGAAAATCCCTTTTCGTCCTTTCTGTATTCG 235
DB 23053 AATTTTATGTGAATAATTAATGATGAAAAAATCTTTCCTACGCTTGTTCTGTATTCG 23112
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 23113 TCCTTACTTACCGCAGCGAAATT 23136

RESULT 11
LOCUS AX044034 349980 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 113 from Patent WO0066791.
ACCESSION AX044034
VERSION AX044034.1 GI:11342918
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1
AUTHORS Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 113 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Location/Qualifiers
source
1..349980
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 bases-seq 108: 300001 to 649980 bases-seq 110:
900001 to 1249980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 bases-seq
113: 1800001 to 2149980 bases-seq 114: 2100001 to
2272325 bases"
BASE COUNT 86473 a 95646 c 85908 g 81953 t
ORIGIN
Query Match 19.5%; Score 50.4; DB 6; Length 349980;
Best Local Similarity 75.0%; Pred. No. 0.00062;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 176 AACTTTTATGTGAACCTCCCATATGAGAAAATCCCTTTTCGTCCTTTCTGTATTCG 235
DB 323053 AATTTTATGTGAATAATTAATGATGAAAAAATCTTTCCTACGCTTGTTCTGTATTCG 323112
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 323113 TCCTTACTTACCGCAGCGAAATT 323136
QY 236 TCCTACTTACTGCCAGCGAAATT 259
DB 323113 TCCTTACTTACCGCAGCGAAATT 323136

RESULT 12
LOCUS AX346615/c 13326 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1686 from Patent WO0200928.
ACCESSION AX346615
VERSION AX346615.1 GI:18494501
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1686 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers

```

```

1. 13326
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT      3846 a 115 c 2666 g 6899 t
ORIGIN

Query Match      15.9%; Score 41.2; DB 6; Length 13326;
Best Local Similarity 54.7%; Pred. No. 0.31;
Matches 82; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 36 TGACTACAAACCAAAACCACTACACCTTAATCACTCACTCAAAACCTCTTCAGA 95
DB 8480 TTATCTATAAACAATAATACATATACCTCCATACCTCCACACACACGACAACTCA 8421
QY 96 CAACCTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTGCTGAAGAATACCGTTGCCG 155
DB 8420 ACATATCTCATCTTCTACAAAAAACTCCAAAAATTTCTTTTACCACTCCCAATTAC 8361
QY 156 AGCCCCAAACCCGTACTGCAACTTTTATT 185
DB 8360 TACCACATAACACATACCTTAAATATAAT 8331

RESULT 13
T2K10/c
LOCUS
DEFINITION
Arabidopsis thaliana chromosome 1 BAC T2K10 sequence, complete
sequence.
ACCESSION
AC005966
VERSION
AC005966.2 GI:12039260
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 90284)
Vysotskaya, V.S., Schwartz, J.R., Yu, G., Toriumi, M., Lenz, C., Liu, S.,
Li, J., Kremenetskaia, I., Luros, J., Gonzalez, A., Altieri, H.,
Araujo, R., Bucher, B., Chao, Q., Conn, L., Conway, A.B., Dunn, P.,
Hansen, N., Huizar, L., Kim, C., Palm, C., Rowley, D., Shinn, P.,
Walker, M., Davis, R.W., Ecker, J.R., Federspiel, N.A. and Theologis, A.
Arabidopsis thaliana chromosome 1 BAC T2K10 sequence
Unpublished
2 (bases 1 to 90284)
Theologis, A.
Direct Submission
Submitted (14-NOV-1998) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
3 (bases 1 to 90284)
Theologis, A.
Direct Submission
Submitted (29-JAN-1999) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
4 (bases 1 to 90284)
Theologis, A.
Direct Submission
Submitted (09-FEB-1999) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
5 (bases 1 to 90284)
Theologis, A.
Direct Submission
Submitted (10-FEB-1999) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
6 (bases 1 to 90284)
Theologis, A.
Direct Submission
Submitted (05-JAN-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
On Jan 5, 2001 this sequence version replaced gi:4204173.
This sequence is of BAC T2K10 from Arabidopsis thaliana chromosome

```

1. In order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide overlap between overlapping submitted clones. The 3' end of this sequence overlaps by 2447 bp to the 5' end of the sequence of BAC T13D8.

Location/Qualifiers

```

1. 90284
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="T2K10"
/complement(<1..871)
/gene="T2K10.1"
/complement(join(<1..199,543..871))
/gene="T2K10.1"
/notes="Similar to gb|AF039182 probable aldo-keto reductase from Fragaria x ananassa. This gene may be cut off. EST gb|U74151 comes from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4487.1"
/db_xref="GI:4249390"
/translation="MSLTITVPTLAIRSGPSHSMVPLGFTAAASPEPTMLKETVI
PAIKLGYRHFDTSPRYQTEPIGALAEAVSLGVRSEFFVTTKLCKADAHGLV
PAIKSLKNLKLVDLYIIHWPVSSKPKYKFFIDEDDFMDFEVVWSEMECQL
GLAKCIGVSNFSCCKL"
1405. 2487
/gene="T2K10.2"
1405. 2487
/gene="T2K10.2"
/notes="Strong similarity to gi|2829864 F316.6 zinc metalloproteinase homolog from Arabidopsis thaliana BAC gb|AC002396. EST gb|226412 comes from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4473.1"
/db_xref="GI:4249376"
/translation="WRTLLTILIFFFTVNPISAKFTVNSIIPPLQFLNATONAMET
FSKLACGHNINGLSKLKQVFRFGYITTTGCTDDFDDVLOSAINYQKNFLKV
TGKLDSSLTROIIVKPRCCNPDLIDVSEMGGKILRTTEKYSFPGKPRWPKRDLT
YAPAFQNNLTREVKEVSEAEATRWNTPLNFTSESLIRADIVIGSFSGHGEPEP
DAMGTLAHSPTGCMHLHGDGDLNISGEISRRILPVTWVLDLSVAVHEIGLL
GDSVSRWGMWIPPAISGDRKVELAKDDIEGIQHLGNGNPGDGGGSKPSRESQSTG
CISVSRWGMWISLSIATCFLISV"
8873. 11295
/gene="T2K10.3"
1018(8873. 9018,9166..9289,9521..9603,9732..9843,
3947..10051,10146..10196,10580..10688,10783..10985,
11074..11295)
/gene="T2K10.3"
/notes="Similar to gi|2829865 F316.4 from Arabidopsis thaliana BAC gb|AC002396."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4474.1"
/db_xref="GI:4249377"
/translation="MATHSSRSNKDAGEDELARRNPYEVLGIPSNSTDGEIKSAYR
RNALRYHPDKNPDDPVAAMFKEVTFAYEVLSDPENRLYDTTSEAVGPNEDLELD
LSSLGAVNTIFAALFKLGVIQITVSANLGEALNGTITPLMVGQVSRKQKQS
AHFYSVTTEEAADGLICKVHSSAKNFKLLYDQVNGGLSLAQALSRDPEITGF
FKRLDAFOPFITEKAGSHVFAVYVDNFKSVSYLTLEIFSSAPFNGEKESLRSTEQ
IVSKETELKFEAEVHEVFAQCTGEMASKCTGEVOELDELLKRRNEICAAVTFPPTQ
GSKRSWSKKSSLLMEPREGEVAVREESGVKKKKWYNTQLRQDKKN"
complement(18173..>20461)
/gene="T2K10.4"
complement(join(18173..18222,18296..18347,18479..18631,
18720..19026,19139..19599,19676..19789,19878..20165,
20369..>20461))
/gene="T2K10.4"
/notes="Identical to gb|AJ010471 mRNA for DEAD box RNA
helicase (RH22) from Arabidopsis thaliana. EST gb|Y11191
comes from this gene."
/codon_start=1

```

FEATURES
source

gene

CDS

gene

CDS

gene

CDS

gene

CDS


```

/evidence=not experimental
/protein_id="AADI4475.1"
/db_xref="GI:4249378"
/translation="DHTVSWKSLGSLDNVSTALRDSQDFRSLTQAVCIPIRSLGKDV
IVAAITGSGKTHGLIADITNTALDSEVTNEERPFPLKNIISLILCPNVLCEQV
VMVNGLVDEGNPLLRVEAVCGSQGPDRLDIIIVSTPAALLNNEIPKNNRLEFLR
CKVYVPEADAMLLCGSFONOIIRLNNMLRDEKOVSLRSLAKNLRGWEIDNASVPOID
LENEDDAFEDSGSISEEDDEEEYLLDIAQSPSEVAEAGSDTKKGRVRKIYTRS
KOYIFATLIPNGKTAGGLKMHFQDAVWVSGNLFHRNSPRUKQKRVETVVDSDVD
ILAEVAKNNNTNERTVFANTVEAVEADILEKASIOCYRHKHKLDERANILA
DRETGVGVCTDAAARQDVNVSHVIQADFASAVDFLHRIQRTARAGQYGTVSL
YTEARDIVAEIRAAVKMGQVETAFSRKGRFRNKVKRAFLKAEAEAEQVAVRY"
complement(21482..22344)
/gene="T2K10.5"
/complement(join(21482..21811,21898..22344))
/gene="T2K10.5"
/notes="Strong similarity to gb|X82030 chloroplast RNA
binding protein (RNPI) from Phaseolus vulgaris."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4476.1"
/db_xref="GI:4249379"
/translation="MAASSVLLSCFKSHQHQLNHLGNYSSNQTLIRFNSTPKPLLI
SSRCSRFRVLSITTVKLEEBEKDDGASAVLDPAAVNTKLYFGNLVYNVSATLA
QIIQDFANPELVLYNRDTCQSGFAPVTMSNVEDCNIIIDNLDGTGYLGRALKNF
ADKPKNPELYPETEHLFVGNLSWTVTSSLAGAFRECGDVVGARVDFDGTGRSR
GYGFVYSSKAEMLTALESLDGFELEGRAIRVNLQAQKPF"
complement(23464..24238)
/gene="T2K10.6"
/complement(join(23464..23568,23729..23859,23953..24238))
/gene="T2K10.6"
/notes="ESTs gb|Z37637, gb|AA042498 and gb|AA042269 come
from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4477.1"
/db_xref="GI:4249380"
/translation="MGKCAQVDAALVLQHPDGKIDRYGVPVSEIMRMYPGHYVSL
IIPKKNIPATITTDKSKRVFRTRVKLRLPTENLVIGHAYRLITQEVYKVLIR
AKPKATKKHOSSTSEKKKPSSEKKIDSESKQNLETDEKQSRVLTNSASSRSKT
WRPSLOISSEATS"
complement(29207..32926)
/gene="T2K10.7"
/complement(join(29207..32031,32167..32926))
/gene="T2K10.7"
/notes="Strong similarity to gb|AF039376 Evelknievel
retrotransposon polyprotein from Arabidopsis arenosa."
/codon_start=1
/evidence=not experimental
/protein_id="AADI4478.1"
/db_xref="GI:4249381"
/translation="MTPTSGSTETIAFSETPTLLNVNMANITKLPTNYIMNRQVH
ALLDGYDLAGVIDGVTAPSMITTAGVSAANPAYKWKQDKLIYSAILGTITTTIO
PLLSRNTAAIEWKLKSIYATPSWGHIOQWROHIKWSKGTITITEYFQGHTRTDE
LALLGKPLEHAQEIEFLGLSEDPYKSVQCTEIRDKPPTLTTELKLINEBAKLWCA
NAATPSYPATAHAANYKGNSSNNNNYNNNNKSHGSSYSSQSSIPASPTPQPPRAN
AATASPNANWLDSDGATHITSDLANLHLQPTGGEDVTIADGSLGSIHTGSGAL
ITDSRLALTDVLYNVNHNKLSIVYRMCNANKSVSEFPFAHQVQKDLXGTQLLQ
RTKDELEWPNYPKPSHFHTTTPKTDLTWSHSLRGHPISLTKLVVVSQFSLPVNS
LQKQFNCSDCLNKKTHLPHTNITSTQPLEYIDLTWSPVSIIDNPKYVILVIDH
YTRYSVYPIKQKSHVDVMTFKALVANKFORKIJHLYSDNGEFTALSFLLSNGI
THLTPHTPHNGISERKPHIVETGTLILQGSMPKSYSWAFIAYILINRMSD
VIGLISYKRLFGVYRPNKLRKVGCLCFWLRPTITTKLDRPAPCFJGYSQSA
YLCNLRITGRTYVGRHVQFVENTYPTFKPLDPTNLEENNNHSITVPSPFPVLP
SVPTTPDPHPQSPAPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
GLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SSNSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
MSBFENAGLRNSTDVLPKPHQNFVTRMFTIKYNPDGSIYERKARFAGKPHQH
GLDYSNTPSPVUKSTTVQTLVDIAVRSWIRQLDINNAFLQRLIEDYVVAOPPFI
NPDRNVVCHLKLHGLKQAPRAWYQELRGFLTCGTINSVNSLIFRQHNKDYIY
ILVYDDFLITKMLDAKTVSTPSPKPLTSGFTIDNPGGYRQILGSL"
complement(41797..44452)

```

```

/gene="T2K10.8"
/complement(join(41797..42015,42114..42274,42362..42531,
42626..42684,42760..42846,42942..43038,43137..43321,
43404..43463,43542..43617,43705..43820,43880..44056,
44142..44239,44326..44452))
/gene="T2K10.8"
/notes="Strong similarity to gi|3337350 F13P17.3 putative
permease from Arabidopsis thaliana BAC gB|AC004481."
Query Match 15.8%; Score 41; DB 8; Length 90284;
Best Local Similarity 56.2%; Pred. No. 0.35;
Matches 77; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 41 TACAAAACCAAAACACCTACCACCTTAATCACTCCACAAACCTCTTCAGACAAC 100
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48143 TAAATATAATAATTAATTCAGTAAAGCAATATAATAATTCACAAACCTTAAGAAGACT 48084
QY 101 TCGTTTTTTTAAAAACAATCTGTAAACAGATAACTGCTGAAGAATACCGTTGCCGAGCCC 160
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48083 GCCTGCATTTATAAATAATTTGTGGCAATTAATTCAGGAAAATACACTTAGCGCTTTC 48024
QY 161 CAAAACCCGTACTGCAA 177
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48023 ATAAAAACATAATAAAA 48007
RESULT 14
LOCUS AC125553 236159 bp DNA linear HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-4H12, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
AC125553
VERSION AC125553.4 GI:30467820
HGTS HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 236159)
AUTHORS Muzny,D,Marie,, Metzker,M.Lee,, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,O., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gubaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kowar,C.,
Kowis,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,M., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,N., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelam,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,

```

Plopper, F., Poindester, A., Popovic, D., Reeves, K., Regier, M.A., Reigh, R., Puazo, M., Quirio, J., Rachlin, E., Rees, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., S., Sanders, W., Savary, G., Scher, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmami, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 236159)
Worley, K.C.

Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236159)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24941144.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCHD
Center clone name: CH230-4H12
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 227359 bases at least Q40
Consensus quality: 228996 bases at least Q30
Consensus quality: 229931 bases at least Q20
Estimated insert size: 238976; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 234738: contig of 234738 bp in length
234838: gap of unknown length
234839 236159: contig of 1321 bp in length.
Location/Qualifiers

1..236159
/organism="Rattus norvegicus"
/mol_type="Genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4H12"
39881..42222
/note="wgs_contig"

BASE COUNT 65808 a 48158 c 47610 g 69047 t 5536 others
ORIGIN

Query Match 15.7%; Score 40.6; DB 2; Length 236159;
Best Local Similarity 52.0%; Pred. No. 0.46; Indels 0;
Matches 91; Conservative 0; Mismatches 84;

QY 50 AAAAAACAACCTTACCAACCTTAACTCAACCTTCCACAAACCTTCTTCAGACAACCTCTTTT 109
Db 171795 ACATATATAGATACAGAGTCAATATATAGAACCAAAATGTATTATCAACATTAATCTTT 171854

QY 110 GAAAAACAATCTGTAAACAGATACCTGTGAAGATACCTTGCAGACCTTCCGAGCCCAAAACCG 169
Db 171855 TAAAAGGAATTTCAACCTGGTGTCTAGTATTATATATCTTCCAGTCCCAACCATCCC 171914

QY 170 TACTGCAACTTTTATGTGAATCTCCCATTTATGAAAAATCCCTTTTCCTCTCT 224
Db 171915 CTCTCAACTCTCTGTGTCTCTCCCTCTCAACCAAGCCCTTTTCATCTAT 171969

RESULT 15
AC126527
LOCUS
DEFINITION
Rattus norvegicus clone CH230-157M5, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC126527
AC126527.5 GI:30579006
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 23253)
Muzzey, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulseghe, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, P., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 167.159 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-92

Perfect score: 259

Sequence: 1 AATTGCTATATCAAGTAGG.....TACTTACTGCCAGCAATT 259

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

```
1: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1980.DAT.*
2: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1981.DAT.*
3: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1982.DAT.*
4: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1983.DAT.*
5: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1984.DAT.*
6: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1985.DAT.*
7: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1986.DAT.*
8: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1987.DAT.*
9: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1988.DAT.*
10: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1989.DAT.*
11: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1990.DAT.*
12: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1991.DAT.*
13: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1992.DAT.*
14: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1993.DAT.*
15: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1994.DAT.*
16: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1995.DAT.*
17: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1996.DAT.*
18: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1997.DAT.*
19: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1998.DAT.*
20: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA1999.DAT.*
21: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA2000.DAT.*
22: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA2001.DAT.*
23: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA2002.DAT.*
24: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA2003.DAT.*
25: /SIDSL1/gcgdata/Geneseq/Geneseq-nuc/NA2004.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	259	100.0	259	19 AAV03600	Neisseria meningitidis
2	77.6	30.0	135	25 ABZ41613	N. gonorrhoeae nuc
3	56.8	21.9	1275	20 AAZ12106	Neisseria meningitidis
4	52.6	20.3	1572	21 AAI15298	DNA encoding a pol
5	52.6	20.3	1575	20 AAZ12107	Neisseria gonorrhoeae
6	52.6	20.3	1575	21 AAI15323	DNA encoding a pol
7	50.4	19.5	12438	21 AAI15500	N. meningitidis pa
8	50.4	19.5	172325	21 AAF21613	Neisseria meningitidis

9	50.4	19.5	349980	21 AAF21612	Neisseria meningitidis
10	50.4	19.5	837096	21 AAI1489	N. meningitidis pa
11	41.2	15.9	13326	24 ABU37113	Human immune system
12	38.2	14.7	15282	24 ABU70189	Chemically treated
13	38.2	14.7	15282	24 AAS61139	Human gene regulat
14	38	14.7	6944	24 ABK34027	Human DNA for stag
15	38	14.7	16509	24 ABL33320	Human immune system
16	37.8	14.6	5552	24 ABL33259	Human immune system
17	37.6	14.5	256	21 AAI1357	N. meningitidis Me
18	37.6	14.5	824	20 AAZ12104	Neisseria meningitidis
19	37.6	14.5	1575	20 AAZ12105	Neisseria meningitidis
20	36.6	14.1	8576	24 ABL34228	Human immune system
21	36.4	14.1	2200	24 ABL34190	Human immune system
22	36.4	14.1	6239	24 ABL70584	Chemically treated
23	36.4	14.1	6239	24 AAS61287	Human gene regulat
24	36.4	14.1	6239	24 ABK31387	Signal transductio
25	36.4	14.1	7316	24 AEN80241	Human chemically m
26	36.4	14.0	8781	24 ABL33687	Human immune system
27	36	13.9	335	22 AAI180630	Human polynucleoti
28	36	13.9	12133	24 AEN80323	Human chemically m
29	35.8	13.8	6099	24 ABL70398	Chemically treated
30	35.8	13.8	6099	24 AAS61346	Human gene regulat
31	35.8	13.8	6099	24 ABK31433	Signal transductio
32	35.8	13.8	7588	24 ABL33103	Human immune system
33	35.6	13.7	3499	21 AAE62298	Caenorhabditis ele
34	35.6	13.7	11996	24 ABL34493	Human metastasis a
35	35.4	13.7	621	24 ABQ46216	Oligonucleotide fo
36	35.4	13.7	621	24 ABQ46217	Oligonucleotide fo
37	35.4	13.7	6534	24 ABL32471	Human immune system
38	35.4	13.7	6906	24 AEN80011	Human chemically m
39	35.4	13.7	10569	24 AEN80264	Human gene regulat
40	35.4	13.7	10647	24 AAS61397	Tumour suppressor
41	35.2	13.6	9007	22 AAS46616	Oligonucleotide fo
42	34.8	13.4	623	24 ABQ43170	Oligonucleotide fo
43	34.8	13.4	623	24 ABQ43171	Oligonucleotide fo
44	34.8	13.4	625	24 ABQ40448	Oligonucleotide fo
45	34.8	13.4	625	24 ABQ40449	Oligonucleotide fo

ALIGNMENTS

```
RESULT 1
AAV03600
ID AAV03600 standard; DNA; 259 BP.
XX
AC AAV03600;
XX
DT 22-OCT-1998 (first entry)
XX
DE Neisseria meningitidis DNA sequence E45.
XX
KW N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW meningitis; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9802547-A2.
XX
PD 22-JAN-1998.
XX
PF 11-JUL-1997; 97WO-FR01295.
XX
PR 12-JUL-1996; 96FR-0008768.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (SMWK ) SMITHKLINE BEECHAM.
XX
PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI Vinals C;
XX
```

DR WPI; 1998-110594/10.
 XX Genes present in *Neisseria meningitidis* but not other *Neisseria*
 PT species - and related host cells, RNA, anti-sense sequences,
 PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*
 PT meningitidis infection and in protective vaccines
 XX
 PS Example 4; Pages 126-127; 150pp; French.
 XX
 CC AAV03575-606 represent sequences that are present in *Neisseria*
 CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the
 CC genes involved in biosynthesis of the capsule polysaccharide, fpaA or C,
 CC opa, rotamase, sequence IC1106, IgA protease, pilin, pilC,
 CC proteins which bind transferrin and opacity proteins. The DNA sequences
 CC are responsible for the differences in pathogenicity between *N.*
 CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that
 CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences
 CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*
 CC *lactamica*, are responsible for colonisation and penetration of the
 CC mucosa. The DNA sequences can be used to produce probes and primers, and
 CC antibodies produced against the encoded proteins are used in standard
 CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*
 CC infection, particularly meningitis.
 XX
 SQ Sequence 259 BP; 81 A; 71 C; 34 G; 73 T; 0 other;
 Query Match 100.0%; Score 259; DB 19; Length 259;
 Best Local Similarity 100.0%; Pred. No. 1.9e-71;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATTGTGTATCAAGTAGGATGGGCAATTTATGCTGACCTACAAAACCAAAACCAACCT 60
 Db 1 AATTGTGTATCAAGTAGGATGGGCAATTTATGCTGACCTACAAAACCAAAACCAACCT 60
 QY 61 ACCACCTTAACTCACTCCACAAACCTCTTTCAGACAACTCGTTTTTGAAGAAACAAATC 120
 Db 61 ACCACCTTAACTCACTCCACAAACCTCTTTCAGACAACTCGTTTTTGAAGAAACAAATC 120
 QY 121 TGTAACAGATAAATGCTGTAAGATAACCGTTGCGAGCCCCCAAAACCGTACTGCAACTT 180
 Db 121 TGTAACAGATAAATGCTGTAAGATAACCGTTGCGAGCCCCCAAAACCGTACTGCAACTT 180
 QY 181 TTATGTGAATTCCTCCATTATGAGAAATCCCTTTTCGTCCTCTTCTGATTGTCCTCT 240
 Db 181 TTATGTGAATTCCTCCATTATGAGAAATCCCTTTTCGTCCTCTTCTGATTGTCCTCT 240
 QY 241 ACTTACTGCCAGCAAAATT 259
 Db 241 ACTTACTGCCAGCAAAATT 259
 RESULT 2
 ABZ41613/c
 ID ABZ41613 standard; DNA; 135 BP.
 AC ABZ41613;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE *N. gonorrhoeae* nucleotide sequence SEQ ID 7815.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 FN WO200279243-A2.
 XX
 FD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB02069.
 XX
 PR 12-FEB-2001; 2001GB-0003424.
 XX

PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Masignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR P-PSDB; ABP0643.
 XX
 PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
 PT medicament for treating or preventing *N. gonorrhoeae* infection
 XX
 PS Disclosure; Page 758; 815pp; English.
 XX
 CC The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
 CC molecules of the invention.
 XX
 SQ Sequence 135 BP; 26 A; 21 C; 38 G; 50 T; 0 other;
 Query Match 30.0%; Score 77.6; DB 25; Length 135;
 Best Local Similarity 95.2%; Pred. No. 1.4e-14;
 Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 80 ACAACACCTCTTCAGACAACTCGTTTTTGAAGAAACAAATCTGTAACAGATAAATCTGCTG 139
 Db 84 ACAACACCTCTTCAGACAACTCGTTTTTGAAGAAACAAATCTGTAACAGATAAATCTGCTG 25
 QY 140 AAGAATACCGTTGCGAGCCCCAA 163
 Db 24 AAGAATACCGTTGCGAGCCCCAA 1
 RESULT 3
 AAZ12106
 ID AAZ12106 standard; DNA; 1275 BP.
 XX
 AC AAZ12106;
 XX
 DT 08-OCT-1999 (first entry)
 XX
 DE *Neisseria meningitidis* strain A complete ORF81 sequence.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.
 XX
 OS *Neisseria meningitidis*.
 XX
 FN WO924578-A2.
 XX
 PD 20-MAY-1999.
 XX
 PF 09-OCT-1998; 98WO-IB01665.
 XX
 PR 01-SEP-1998; 98GB-0019016.
 PR 06-NOV-1997; 97GB-0023516.
 PR 14-NOV-1997; 97GB-0024190.
 PR 18-NOV-1997; 97GB-0024386.
 PR 27-NOV-1997; 97GB-0025158.
 PR 10-DEC-1997; 97GB-0026147.
 PR 14-JAN-1998; 98GB-0000759.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
 XX
 DR WPI; 1999-327407/27.
 DR P-PSDB; AAY38652.
 XX
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for

PT diagnosis, treatment and prevention of infection
 PS Claim 9; Page 206; 524pp; English.

XX Nucleotide sequences AAZ11972-Z12358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.

XX SQ Sequence 1275 BP; 300 A; 304 C; 319 G; 352 T; 0 other;

Query Match 21.9%; Score 56.8; DB 20; Length 1275;
 Best Local Similarity 96.7%; Pred. No. 1e-07;
 Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 200 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 259
 Db 1 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 60

RESULT 4
 AAA15298
 ID AAA15298 standard; DNA; 1572 BP.

XX AC AAA15298;

XX DT 04-SEP-2000 (first entry)

XX DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.

XX KW Pathogenic strain: *Neisseria*; vaccine; *Neisseria* infection; ss.

XX OS *Neisseria meningitidis*.

XX FH Key Location/Qualifiers
 FT 1..1572
 FT CDS /*tag= a

XX PN WO200026375-A2.

XX PD 11-MAY-2000.

XX PF 28-OCT-1999; 99WO-FR02643.

XX PR 30-OCT-1998; 98FR-0013693.

XX PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Aujame L, Bouchardon A, Renaud-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;

XX DR WPI; 2000-365622/31.
 DR P-PSDB; AAY93267.

XX PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 PT or preventative vaccines and for diagnosis

XX PS Claim 2; Page 56-59; 187pp; French.

XX The present sequence encodes a protein that is specific for pathogenic
 CC strains of *Neisseria*. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.

XX SQ Sequence 1572 BP; 380 A; 383 C; 395 G; 414 T; 0 other;

Query Match 20.3%; Score 52.6; DB 21; Length 1572;
 Best Local Similarity 93.2%; Pred. No. 2.3e-06;
 Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 200 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 258
 Db 1 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 59

RESULT 5
 AAZ12107
 ID AAZ12107 standard; DNA; 1575 BP.

XX AC AAZ12107;

XX DT 08-OCT-1999 (first entry)

XX DE *Neisseria gonorrhoeae* complete ORF81 sequence.

XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW treatment; *Neisseria* infection; meningitis; septicaemia; gonorrhea; ss.

XX OS *Neisseria gonorrhoeae*.

XX PN WO9924578-A2.

XX PD 20-MAY-1999.

XX PF 09-OCT-1998; 98WO-IB01665.

XX PR 01-SEP-1998; 98GE-0019016.

XX PR 06-NOV-1997; 97GB-0023516.

XX PR 14-NOV-1997; 97GB-0024190.

XX PR 18-NOV-1997; 97GB-0024386.

XX PR 27-NOV-1997; 97GB-0025158.

XX PR 10-DEC-1997; 97GB-0026147.

XX PR 14-JAN-1998; 98GE-0000759.

XX PA (CHIR-) CHIRON SPA.

XX PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;

XX DR WPI; 1999-327407/27.

XX DR P-PSDB; AAY38653.

XX PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection
 XX PS Claim 9; Page 208; 524pp; English.

XX CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
 CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
 CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
 CC fragments, their nucleic acids and antibodies are used for diagnosis,
 CC prevention (as vaccines) or treatment of *Neisseria* infections,
 CC such as meningitis, septicaemia and gonorrhea. Both organisms
 CC are closely related. Fragments of the nucleic acids are useful
 CC as hybridisation probes and antisense reagents.

XX SQ Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;
 Query Match 20.3%; Score 52.6; DB 20; Length 1575;
 Best Local Similarity 93.2%; Pred. No. 2.3e-06;
 Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 200 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 258
 Db 1 ATGAGAAATCCCTTTTCGTCCTCTTTCTGTTATTCGTCCTACTTACTGCCAGCGAAAT 59

RESULT 6
 AAA15323

```

ID  AAA15323 standard; DNA; 1575 BP.
XX
AC  AAA15323;
XX
DT  04-SEP-2000 (first entry)
XX
DE  DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
KW  Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS  Neisseria gonorrhoeae.
XX
FH  Key Location/Qualifiers
FT  CDS 1..1575
FT  /*tag= a
XX
PN  WO200026375-A2.
XX
PD  11-MAY-2000.
XX
PF  28-OCT-1999; 99WO-FR02643.
XX
PR  30-OCT-1998; 98FR-0013693.
XX
PA  (INRM ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA  (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI  Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI  Tinsley C, Perrin A;
XX
DR  WPI; 2000-365622/31.
DR  P-PSDB; AAY93293.
XX
PT  New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT  or preventative vaccines and for diagnosis
XX
PS  Claim 4; Page 137-139; 187pp; French.
XX
CC  The present sequence encodes a protein that is specific for pathogenic
CC  strains of Neisseria. The polynucleotides, polypeptides, or their
CC  antigenic fragments, are used in vaccines to treat or protect against
CC  Neisseria infections, particularly by N. meningitidis. The
CC  polynucleotide sequence is also used for recombinant production of
CC  the polypeptide and to produce attenuated Neisseria strains that
CC  overexpress it, or express it in a non-toxic mutant form.
XX
SQ  Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 other;

Query Match 20.3%; Score 52.6; DB 21; Length 1575;
Best Local Similarity 93.2%; Pred. No. 2.3e-06;
Matches 55; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 200 ATGAGAAATCCCTTTTCGCTCTTCTGTATTGTCCTACTTACTGCGCAGGAAT 258
Db 1 ATGAGAAATCCCTTTTCGCTCTTCTGTATTGTCCTACTTACTGCGCAGGAAT 59

RESULT 7
AAA81550
ID AAA81550 standard; DNA; 12438 BP.
XX
AC AAA81550;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_97 SEQ ID NO:97.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX

```

```

PN  WO200022430-A2.
XX
PD  20-APR-2000.
XX
PF  08-OCT-1999; 99WO-US23573.
XX
PR  09-OCT-1998; 98US-0103794.
PR  30-APR-1999; 99US-0132068.
XX
PA  (CHIR ) CHIRON CORP.
XX
PI  Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI  Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI  Rappuoli R, Fizza M;
XX
DR  WPI; 2000-318079/27.
XX
PT  Isolated nucleotide sequences of Neisseria meningitidis which can be
PT  used in the diagnosis and treatment of N. meningitidis infection and
PT  other Neisserial infections, for example, N.gonorrhoea
XX
PS  Claim 7; Page 1489-1492; 1760pp; English.
XX
CC  The present invention describes methods of obtaining immunogenic
CC  proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC  represent specifically claimed Neisseria meningitidis genomic DNA
CC  sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC  Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC  AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC  isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC  AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC  sequences, which are all used in the exemplification of the present
CC  invention. The nucleic acid sequences, protein sequences and antibodies
CC  against them, can be used in the manufacture of a composition. The
CC  composition can be used as a medicament (or in the manufacture of a
CC  medicament) for treating, preventing or diagnosing infection due to
CC  Neisserial bacteria. For example, some of the identified proteins could
CC  be components of vaccines against Meningococcus B; against all serotypes;
CC  and/or against all pathogenic Neisseriae. Identification of sequences
CC  from the bacterium will also facilitate production of biological probes,
CC  particularly organism-specific probes. Attempts to make efficacious
CC  Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC  Multivalent vaccines have also been tried but none have successfully
CC  overcome antigenic variability. The provision of further, complete
CC  sequences may provide an opportunity to identify secreted or surface
CC  exposed proteins that may be presumed targets for the immune system and
CC  which are not antigenically variable or at least more conserved than
CC  other more variable regions.

Query Match 19.5%; Score 50.4; DB 21; Length 12438;
Best Local Similarity 75.0%; Pred. No. 2.3e-05;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 176 AACTTTTATGTGAACCTCCCATTTATGAGAAAACCCCTTTTCGCTCTTCTGTATTGCG 235
Db 2873 AATTTTATGTGAATATTATGATGAAAAAATCTTCTTACGCTTGTCTGATTGCG 2932

QY 236 TCCCTACTTACTGCCAGCGAAATT 259
Db 2933 TCTTTACTTACGCCAGCGAAATT 2956

RESULT 8
AAF21613
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.

```



```

XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX Neisseria meningitidis.
OS WO200065791-A1.
XX 09-NOV-2000.
XX 08-MAR-2000; 2000WO-US05928.
XX 30-APR-1999; 99US-0132068.
XX 08-OCT-1999; 99WO-US23573.
XX 28-FEB-2000; 2000GB-0004695.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CW, Grandi G;
XX WPI; 2000-647603/62.
XX Neisseria meningitidis B full length genome sequence and open reading
frames are used to detect, treat and prevent Neisserial infections -
XX Claim 7; Appendix A; 692pp; English.
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;
Query Match 19.5%; Score 50.4; DB 21; Length 172325;
Best Local Similarity 75.0%; Pred. No. 5.7e-05;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 176 AACCTTTTATGTGAACTCCCATATGAGAAATCCCTTTTCGCTCTTCTGTATTCG 235
DB 23053 AATTTTATGTGAAATATTAATGATGAAATACTTTCCTACGCTGTCTGTATTCG 23112
QY 236 TCCCTACTACTGCCAGCAAAATT 259
DB 23113 TCTTACTACCGCAGCAAAATT 23136
RESULT 9
AAF21612
ID AAF21612 standard; DNA; 349980 BP.
XX

```

```

AC AAF21612;
XX 13-MAR-2001 (first entry)
XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX Neisseria meningitidis.
OS WO200065791-A1.
XX 09-NOV-2000.
XX 08-MAR-2000; 2000WO-US05928.
XX 30-APR-1999; 99US-0132068.
XX 08-OCT-1999; 99WO-US23573.
XX 28-FEB-2000; 2000GB-0004695.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CW, Grandi G;
XX WPI; 2000-647603/62.
XX Neisseria meningitidis B full length genome sequence and open reading
frames are used to detect, treat and prevent Neisserial infections -
XX Claim 7; Appendix A; 692pp; English.
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;
Query Match 19.5%; Score 50.4; DB 21; Length 349980;
Best Local Similarity 75.0%; Pred. No. 7.2e-05;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 176 AACCTTTTATGTGAACTCCCATATGAGAAATCCCTTTTCGCTCTTCTGTATTCG 235
DB 23053 AATTTTATGTGAAATATTAATGATGAAATACTTTCCTACGCTGTCTGTATTCG 323112
QY 236 TCCCTACTACTGCCAGCAAAATT 259
DB 323113 TCTTACTACCGCAGCAAAATT 323136

```

```
RESULT 10
AAA81489
ID AAA81489 standard; DNA; 837096 BP.
XX
AC AAA81489;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 629-865; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;
XX
Query Match 19.5%; Score 50.4; DB 21; Length 837096;
Best Local Similarity 75.0%; Pred. No. 9.7e-05;
Matches 63; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
XX
QY 176 AACTTTATTGGAACCTCCCATATGAGAAATCCCTTTTCGTCCTCTTCTGATTGCG 235
```

```

Db 687824 AATTTTATTGGAATATTAAATGATGAAAAAATCTTTCCTACGCTGTCTGATTGCG 687883
QY 236 TCCCTACTTACTGCCAGCGAAATT 259
AC ABL33713;
Db 687884 TCTTTACTTACGCCAGCGAAATT 687907

RESULT 11
ABL33713/c
ID ABL33713 standard; DNA; 13326 BP.
XX
XX ABL33713;
AC
XX 26-MAR-2002 (first entry)
DT
XX
DE Human immune system associated gene SEQ ID NO: 1686.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antianaemic; cytostatic; neotropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX neurotrophic; antiarthritic; antidiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 1686; 32pp + Sequence Listing; German.
PS
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
XX Sequence 13326 BP; 3846 A; 115 C; 2656 G; 6699 T; 0 other;
SQ
Query Match 15.9%; Score 41.2; DB 24; Length 13326;
Best Local Similarity 54.7%; Pred. No. 0.018; 68; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
XX
QY 36 TGACCTACAAAACCAAAACCACTTACCACCTTAACTCCACAAACCCCTTTTCAGA 95
Db 8480 TTATCTATAAACAATAACATATACCTCCCATCTCCACACACACAGCAACTCA 8421
QY 96 CAACTCTGTTTTTGAAGAACAACTGTGAACAGATACTCTGAGAAATACCGTTGCGG 155
Db 8420 ACATATCTCATCTCTACAAAACCTCCAAAATTTCTTTTCCACCACTCCCAATTAC 8361
```

QY	156	AGCCCCAAACCGTACTGCACTTTATT	185
Db	8360	TACCACATACACATACCTTAAATATAAT	8331
Db	7932	CGAAACTAAATCTATACCCCAACCTACACCTCCATCTACAAAAATATCCACTTTAAA	7877
QY	220	CCTCTTTCTGTATTGTCCTTAC	242
Db	7872	ATACAAAAATAATTCTCCCTAC	7850
RESULT 12			
ABL70189/c			
ID	ABL70189 standard; DNA; 15282 BP.		
AC	ABL70189;		
XX	01-JUL-2002 (first entry)		
DE	Chemically treated cell signalling DNA sequence#40.		
XX	Cell signalling; cytosine methylation; cell signalling disease;		
KW	cancer; tumour; cytostatic; ds.		
XX	Unidentified.		
OS	WO200202807-A2.		
PN	10-JAN-2002.		
XX	29-JUN-2001; 2001WO-EP07471.		
PF	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043926.		
XX	(EPIG-) EPIGENOMICS AG.		
PA	Olek A, Piepenbrock C, Berlin K;		
PI	WPI; 2002-154758/20.		
XX	Nucleic acid, useful for diagnosis and therapy of diseases associated		
PT	with cell signalling e.g. cancer, comprises chemically modified genomic		
PT	sequences of genes associated with cell signalling		
XX	Claim 1; SEQ ID NO 79; 24pp-sequence listing; English.		
PS	The invention relates to a nucleic acid comprising a sequence of at least		
XX	18 bases of a segment of chemically pretreated DNA of genes associated		
CC	with cell signalling. The activity of the modified sequences of the		
CC	invention may be described as cytostatic. The object of the invention is		
CC	to provide the chemically modified DNA of genes associated with cell		
CC	signalling, as well as oligonucleotides and/or pNA-oligonucleotides for		
CC	detecting cytosine methylations, as well as a method which is		
CC	particularly suitable for the diagnosis and/or therapy of genetic and		
CC	epigenetic parameters of genes associated with cell signalling. The		
CC	chemically modified DNA provided by the invention is useful for diagnosis		
CC	and therapy of diseases such as solid tumours and cancer. The sequences		
CC	given in records ABL70111-ABL70626 represent chemically pre-treated		
CC	genomic DNA's of genes associated with cell signalling.		
CC	Note: The sequence data for this patent is not represented in the printed		
CC	specification, but is based on sequence information supplied by the		
CC	European Patent Office.		
XX	Sequence 15282 BP; 3617 A; 581 C; 4050 G; 7034 T; 0 other;		
SQ	Query Match 14.7%; Score 38.2; DB 24; Length 15282;		
	Best Local Similarity 49.3%; Pred No. 0.17;		
	Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;		
QY	40	CTACAAAAACAAACAACTACACCTTACACCTTAACTAACCTCCAAACCTCTTCAGACAAC	99
Db	8052	CAACAAACAAACAACTAAAAAATCCCCACCTCTCAATTTAAATCTACTTCCAAAAAC	7993
QY	100	CTCGTTTTTGAACACATCTGTAACAGATTAAGTCTGAAGATACCGTTCGCGAGCC	159
Db	7992	AAACTCCCTTTAAACAAAAAACTAACTTCTAAATAAAAAAACTATTTAAAAAAAC	7933
QY	160	CCAAAAACCCGTACTGCACTTTTATGTGAATCTCCCAATATGAAAAATCCCTTTTCGT	219

QY	156	AGCCCCAAACCGTACTGCACTTTATT	185
Db	8360	TACCACATACACATACCTTAAATATAAT	8331
Db	7932	CGAAACTAAATCTATACCCCAACCTACACCTCCATCTACAAAAATATCCACTTTAAA	7877
QY	220	CCTCTTTCTGTATTGTCCTTAC	242
Db	7872	ATACAAAAATAATTCTCCCTAC	7850
RESULT 12			
ABL70189/c			
ID	ABL70189 standard; DNA; 15282 BP.		
AC	ABL70189;		
XX	01-JUL-2002 (first entry)		
DE	Chemically treated cell signalling DNA sequence#40.		
XX	Cell signalling; cytosine methylation; cell signalling disease;		
KW	cancer; tumour; cytostatic; ds.		
XX	Unidentified.		
OS	WO200202807-A2.		
PN	10-JAN-2002.		
XX	29-JUN-2001; 2001WO-EP07471.		
PF	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043926.		
XX	(EPIG-) EPIGENOMICS AG.		
PA	Olek A, Piepenbrock C, Berlin K;		
PI	WPI; 2002-154758/20.		
DR	Nucleic acid, useful for diagnosis and therapy of diseases associated		
PT	with cell signalling e.g. cancer, comprises chemically modified genomic		
PT	sequences of genes associated with cell signalling		
XX	Claim 1; SEQ ID NO 79; 24pp-sequence listing; English.		
PS	The invention relates to a nucleic acid comprising a sequence of at least		
XX	18 bases of a segment of chemically pretreated DNA of genes associated		
CC	with cell signalling. The activity of the modified sequences of the		
CC	invention may be described as cytostatic. The object of the invention is		
CC	to provide the chemically modified DNA of genes associated with cell		
CC	signalling, as well as oligonucleotides and/or pNA-oligonucleotides for		
CC	detecting cytosine methylations, as well as a method which is		
CC	particularly suitable for the diagnosis and/or therapy of genetic and		
CC	epigenetic parameters of genes associated with cell signalling. The		
CC	chemically modified DNA provided by the invention is useful for diagnosis		
CC	and therapy of diseases such as solid tumours and cancer. The sequences		
CC	given in records ABL70111-ABL70626 represent chemically pre-treated		
CC	genomic DNA's of genes associated with cell signalling.		
CC	Note: The sequence data for this patent is not represented in the printed		
CC	specification, but is based on sequence information supplied by the		
CC	European Patent Office.		
XX	Sequence 15282 BP; 3617 A; 581 C; 4050 G; 7034 T; 0 other;		
SQ	Query Match 14.7%; Score 38.2; DB 24; Length 15282;		
	Best Local Similarity 49.3%; Pred No. 0.17;		
	Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;		
QY	40	CTACAAAAACAAACAACTACACCTTACACCTTAACTAACCTCCAAACCTCTTCAGACAAC	99
Db	8052	CAACAAACAAACAACTAAAAAATCCCCACCTCTCAATTTAAATCTACTTCCAAAAAC	7993
QY	100	CTCGTTTTTGAACACATCTGTAACAGATTAAGTCTGAAGATACCGTTCGCGAGCC	159
Db	7992	AAACTCCCTTTAAACAAAAAACTAACTTCTAAATAAAAAAACTATTTAAAAAAAC	7933
QY	160	CCAAAAACCCGTACTGCACTTTTATGTGAACTTCCCAATTATGAAAAATCCCTTTTCGT	219

```

XX SQ Sequence 15282 BP; 3617 A; 581 C; 4050 G; 7034 T; 0 other;
Query Match 14.7%; Score 38.2; DB 24; Length 15282;
Best Local Similarity 49.3%; Pred. No. 0.17;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 40 CTCARAAACCAAAACAACTACACACCTTAATCACTCCACAAACCTCTCAGACAC 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8052 CAACAAACAAACCTAAATAATTCCTCCACCTCTCAATTTTAAATATCTTCCAAAAC 7993
QY 100 CTCCTTTTTGAAAAACAATCTGTAAACAGATAACTGCTGAGAAATACCGTTCGCCAGCC 159
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7992 AAATCCCTTTAAACAAACAAACAACTTCTTAACATAAAAAAACTATTCTAAAAAAAC 7933
QY 160 CCAAAACCGCTACTGCAACTTTTATGTGAACCTCCATATGAGAAATCCCTTTTCGT 219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7932 CGAACTAAATCTATACCCCAACCTACACACCTCCATCTACAAAAATATCCACITTTAA 7873
QY 220 CCTCTTCTGTATTCGCTCCCTAC 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7872 ATACAAAAAATATTCTCCCTAC 7850

RESULT 14
ABK34027/c
ID ABK34027 standard; DNA; 6944 BP.
XX AC ABK34027;
XX DT 18-JUN-2002 (first entry)
XX DE Human DNA for staging of Astrocytomas, complement, #57.
XX KW Human; ds; astrocytoma; cytostatic; staging; cysteine methylation; CpG;
XX KW bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
XX KW matrix assisted laser desorption/ionization mass spectrometry.
XX OS Homo sapiens.
XX WO200202808-A2.
XX 10-JAN-2002.
XX 02-JUL-2001; 2001WO-EP07538.
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-171649/22.
XX PT Novel chemically modified genomic DNA sequences, useful in the
XX PT characterisation, classification, differentiation, grading, staging,
XX PT treatment and/or diagnosis of astrocytomas or predisposition to
XX PT astrocytomas.
XX Claim 1; SEQ ID No 114; 37pp; English.
XX CC The invention relates to a nucleic acid comprising a sequence (I) of at
XX CC least 18 bases in length of a segment of chemically pre-treated genomic
XX CC DNA which has any one of the sequences of (ABK33919-ABK34032) or its
XX CC complement. Also included are an oligonucleotide or peptide nucleic
XX CC acid (or set thereof) of at least 9 nucleotides which hybridises to (I),
XX CC primers for (II), probes for detecting cytosine methylation or single-
XX CC nucleotide polymorphisms (SNP) in (I), an array of oligomers
XX CC or peptide nucleic acids for analysing diseases associated with the
XX CC methylation states of the CpG dinucleotides of (I). The array is useful
XX CC for determining genetic and/or epigenetic parameters, classification,
XX CC differentiation, grading, staging, treatment and/or diagnosis of

```

```

CC astrocytomas, or the predisposition to astrocytomas by analysing cytosine
CC methylations, involves obtaining a biological sample containing genomic
CC DNA, extracting the genomic DNA, converting cytosine bases which are
CC unmethylated at the 5-position, in the genomic DNA sample, to uracil or
CC another base which is dissimilar to cytosine in terms of hybridisation
CC behaviour, by chemical treatment and amplifying chemically pre-treated
CC genomic DNA fragments using the array and a polymerase, where the
CC amplicates carry a detectable label. The method further involves
CC identifying methylation status of one or more cytosine positions, and
CC analysing methylation status of the cytosine positions by reference to
CC one or more data sets. The genomic DNA is chemically treated by using a
CC bisulphite, hydrogen sulphite or disulphite. The amplification
CC step amplifies DNA which is of particular interest in astrocytoma or
CC brain tissue, based on the specific genomic methylation status of brain
CC tissues, as opposed to background DNA. The amplicates carry a
CC fluorescent label or radionuclide. Optionally, the labels of the
CC amplicates are detachable molecule fragments having a typical mass
CC which are detected in a mass spectrometer. The fragments of chemically
CC pre-treated genomic DNA to be amplified, have a single positive or
CC negative charge for a better detectability in the mass spectrometer.
CC Preferably, the amplicates or fragments of the amplicates are
CC detected by matrix assisted laser desorption/ionization mass spectrometry
CC (MALDI) or using electron spray mass spectrometry (ESI). The
CC present sequence is one of the chemically pre-treated reference DNA
CC samples of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 6944 BP; 1633 A; 173 C; 1692 G; 3446 T; 0 other;
Query Match 14.7%; Score 38; DB 24; Length 6944;
Best Local Similarity 54.2%; Pred. No. 0.15;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

```

QY 44 AAACCAAAACAACTACACACCTTAACTCAACTCCACAAACCTCTTCAGACACCTCG 103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4871 AAATCCAAAAAAACAAACGACCTCTATAAAAAAATAAAAAACCGCCCGTAAACAGCA 4812
QY 104 TTTTGTAAAAACAATCTGTAAACAGATAACTGCTGAAGAATACCGTTGCCGAGCCCAA 163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4811 AAATCAAAAAACCGCTCCGAAAAAATACTAAGAAATCTCAACCCCTCTCGCCCCCA 4752
QY 184 AACCGTACTCCAACTTTTATT 185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4751 AACTCCCACTCTTAAAAATATT 4730

```

```

RESULT 15
ABL33320/c
ID ABL33320 standard; DNA; 16509 BP.
XX AC ABL33320;
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 1293.
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;
XX KW antiarteriosclerotic; antianaemic; cytostatic; neurotropic;
XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX KW gene; ds.
XX OS Homo sapiens.
XX WO200200928-A2.
XX 03-JAN-2002.

```

```
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 1293; 32pp + Sequence Listing; German.
PS
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 16509 BP; 3851 A; 379 C; 4791 G; 7488 T; 0 other;
Query Match 14.7%; Score 38; DB 24; Length 16509;
Best Local Similarity 53.1%; Pred. No. 0.2; Mismatches 90; Indels 1; Gaps 1;
Matches 103; Conservative 0;
QY 41 TACAAACCAAAACAACTACACCCCTTAATCACTCCAC-AAACCCCTCTTCAGACAAC 99
Db 11520 TAAATATACAACTAACTACCCCACTCTAAATCGCCACAAACCCCTTTATAACATA 11461
QY 100 CTCGTTTTTGNAAACATCTGTAACAGATACTGCTGAAGATACCGTTGCGAGCC 159
Db 11460 CTTATCAACAAACAAACAAATATAAAATAAAATACTTTAAATATACGAAACCTAAAC 11401
QY 160 CCAAAACCCGCTACTGCACTTTTATGTGAACCTCCCATTTATGAGAAATCCCTTTTGGT 219
Db 11400 TAAATCCGTTTCAATCTTTTACTAATATTCGACCTAAACTAATTTATCCCTAAACT 11341
QY 220 CCTCTTCTGTATT 233
Db 11340 TTTTCTTTTAAAT 11327
```

Search completed: November 15, 2003, 00:35:44
Job time : 169.159 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1727.81 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-92

Perfect score: 259

Sequence: 1 AATGTGTATATCAAGTAGG.....TACTTACTGCCAGCGAAATT 259

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_hic:**

9: gb_est1:**

10: gb_est2:**

11: gb_hic:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_pbg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	45.2	17.5	1101	29	CNS00LT2
C 2	43.2	16.7	1201	13	EX419736
C 3	42.8	16.5	1181	13	EX363791
C 4	42	16.2	1101	29	CNS00DT7

C	5	40.6	15.7	581	12	BM962443
	6	39.6	15.3	1181	13	EX438271
	7	39.4	15.2	1078	28	EX438271
	8	39.2	15.1	939	29	CNS00CNG
C	9	38.6	14.9	712	13	EX416727
C	10	38.4	14.8	1201	13	EX355642
C	11	38.4	14.8	1201	13	EX355642
C	12	38.2	14.7	1091	13	EX424950
C	13	38.2	14.7	1101	29	CNS0182P
C	14	38	14.7	1173	13	EX456623
C	15	37.8	14.6	1201	13	EX338623
C	16	37.8	14.6	1124	13	EX436282
C	17	37.6	14.5	948	9	AL583463
C	18	37.6	14.5	1201	13	EX464312
C	19	37.4	14.4	1101	29	CNS0172T
C	20	37.2	14.4	1196	13	EX42034
C	21	37.2	14.4	2279	11	AK086359
C	22	37	14.3	623	13	BQ831601
C	23	37	14.3	1201	13	EX428660
C	24	36.8	14.2	619	28	AZ282736
C	25	36.8	14.2	926	29	CNS008LM
C	26	36.8	14.2	1186	13	EX403722
C	27	36.8	14.2	1201	9	AL548228
C	28	36.8	14.2	1201	13	EX361080
C	29	36.6	14.1	1146	29	CNS021G2
C	30	36.4	14.1	588	28	AZ455602
C	31	36.4	14.1	1101	29	CNS014NP
C	32	36.4	14.1	1201	13	EX405691
C	33	36.2	14.0	538	9	AW117910
C	34	36.2	14.0	850	29	AG126628
C	35	36	13.9	652	28	AZ784334
C	36	36	13.9	859	29	CNS004YY
C	37	36	13.9	1072	13	EX383295
C	38	36	13.9	1111	13	EX338659
C	39	35.8	13.8	374	9	AW455109
C	40	35.8	13.8	1101	29	CNS0006J
C	41	35.8	13.8	1113	13	EX463898
C	42	35.8	13.8	1180	13	EX463188
C	43	35.8	13.8	1201	9	AL563696
C	44	35.8	13.8	1201	13	EX354226
C	45	35.6	13.7	255	12	BM958531

ALIGNMENTS

RESULT 1
CNS00LT2/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPL1-98 library from Drosophila melanogaster (fruit fly), Genomic survey sequence.
ACCESSION
AL078714
VERSION
AL078714.1 GI:5102004
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Direct Submission
Genoscope.
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Db	650	WYTTAMTCMTWCYCMARAWTATMTCTYMTHTCCYCYCWYVYCTCTCYCTCYVYVYVYT	709
Qy	186	GTGAACCTTCGCAATATGAGAAAAATCCCTTTTCGTCTCTCTTCTGTATTCTGTCCTACTTA	2451
Db	710	YHTTCTTACCTTYMTATATAWCATCTCTTTYYNNNNNNNCCYCYCYCYVYVYVYVY	7699
Qy	246	CT 247	
Db	770	YX 771	

RESULT 5
 BM962443/c , BM962443 581 bp mRNA linear EST 18-MAR-2002
 LOCUS
 DEFINITION
 cihA6K2T Ascidian hemocytes cDNA library Ciona intestinalis cDNA,
 mRNA sequence.
 ACCESSION
 VERSION
 BM962443
 BM962443.1 GI:19545870
 EST.
 KEYWORDS
 SOURCE
 Ciona intestinalis
 ORGANISM
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

CONTACT: Masaharu Satake
Department of Molecular Immunology
Institute of Development, Aging and Cancer, Tohoku University
Seiryō-machi 4-1, Aoba-ku, Sendai 980-8575, Japan
Tel: 81-22-717-8477
Fax: 81-22-717-8482

FEATURES
Location/Qualifiers
Email: satake@idac.tohoku.ac.jp

	Conserved Matches	Identity 76%	Conservative 0	Mismatches 0	Indels 59	Gaps 0
Qy	88	TCCTCAGACAACCTCGCTTTTTCAGAAAACAATCTGTAAACAGATAACTGCTGAAGAATAC	147			
Db	380	TCATAAATAAACAAGCTTTTTCAGAAATTAACCTTTAATAGATAACCGCTACAACACAG	321			
Qy	148	CGTTGCCGAGCCCCAAAACCCGCTACTGCAACTTTTATTGTGAACTTCCCATTTATGAGAAA	207			
Db	320	TTTTGCTGAACCTTAAAAGAAGTAGATTATTGTTTTTAAAGTCATTTTATATCTAG	261			
Qy	208	ATCCCTTTTTCGCTCT	222			
Db	260	ATCTCATCTGCTGCT	246			

[illegible]

SOURCE Homo sapiens (human)

and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

```

source
1..939
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR28H16"
/clone="end : RFI-98"
/note="end : TET3"
71 a 349 c 104 g 180 t 235 others
BASE COUNT
ORIGIN
Query Match 15.1%; Score 39.2; DB 29; Length 939;
Best Local Similarity 11.0%; Pred. No. 4.2;
Matches 22; Conservative 108; Mismatches 70; Indels 0; Gaps 0;

QY 58 CTTACACCCCTTAATCAACTCCAAACCCCTCTTCAGACAACTCGTTTTTGAAACAA 117
Db 353 CTTCCMMYHTTYMCCCCCCCMCMCHTCCMMYMMYMMYMMYMMYMMYMMYMMYMM 412
QY 118 ATCTGTAACAGATACTGCTGAAGATAACCGTTGCCGAGCCCAACCGTACTGCAA 177
Db 413 MMTMMHMMYMTTMMHMTMCMYMTTMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 472
QY 178 CTTTATTGTGAACCTCCATTATGAGAAATCCCTTTTCCTCTTCTTCTGTTATCGTC 237
Db 473 MTTMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 532
QY 238 CTTACTTACTGCCAGCGAAA 257
Db 533 MMTMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 552

```

RESULT 9

```

BX416727/c
LOCUS
DEFINITION
BX416727 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CS0DA011Y114 5-PRIME, mRNA sequence.
ACCESSION
BX416727
VERSION
BX416727.1 GI:30765629
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DA011BE07Q1.
Location/Qualifiers
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA011Y114"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
154 a 53 c 50 g 148 t 307 others
BASE COUNT

```

ORIGIN

```

Query Match 14.8%; Score 38.6; DB 13; Length 712;
Best Local Similarity 3.5%; Pred. No. 5.9;
Matches 5; Conservative 96; Mismatches 40; Indels 0; Gaps 0;

QY 38 ACCTACAAACCAAAACAACTTACCACCTTAATCAACTCCACAAACCTCTTCAGACA 97
Db 612 MMTMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 553
QY 98 ACCTCGTTTTTTGAAAAACAATCTGTAAACAGATACTCTGAAGATAACCGTTGCCGAG 157
Db 552 MMTMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 493
QY 158 CCCCAAAACCCGTACTGCAAC 178
Db 492 CMCMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 472

RESULT 10
BX355642/c
LOCUS
DEFINITION
BX355642 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1002Y015 5-PRIME, mRNA sequence.
ACCESSION
BX355642
VERSION
BX355642.1 GI:30379989
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1002AH08Q1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1002Y015"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
408 a 117 c 179 g 272 t 225 others
BASE COUNT
ORIGIN
Query Match 14.8%; Score 38.4; DB 13; Length 1201;
Best Local Similarity 22.1%; Pred. No. 7.1;
Matches 42; Conservative 77; Mismatches 71; Indels 0; Gaps 0;

QY 44 AAAACCAAAACAACTTACCACCTTAACTCACTCCACAAACCTCTTCAGACAACCTCG 103
Db 1200 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1141
QY 104 TTTTGTAAAAACAATCTGTAAACAGATACTGCTGAAGATAACCGTTGCCGAGCCCA 163
Db 1140 HMTTTHMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 1081
QY 164 AACCCGTACTGCAACTTTTATTGTGAACCTCCCATATGAGAAAATCCCTTTTCGCTC 223
Db 1080 AMMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 1021

```


Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaut at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelO8AC11.

FEATURES
Source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37D10"
/clone_lib="DrosBAC"
/plasmid="pBelO8AC11"
/notes="end : SP6"

BASE COUNT 274 a 268 c 128 g 73 t 358 others
ORIGIN

Query Match 14.7% Score 38.2; DB 29; Length 1101;
Best Local Similarity 14.2%; Pred.No.8;
Matches 31; Conservative 116; Mismatches 72; Indels 0; Gaps 0;
QY 26 CATTATGCTGCTGACTCAAAACCAAAACACCTACGACCCCTTAATCAACTCCACAAC 85
Db 1097 HHHHWAHVMHHCHVMHYMMHYMWHCHYMAHAYMWHMTAYAHMCHAHVYTW 1038
QY 86 CCTCTTCAGACAACCTCGTTTTTGA AAAACAATCTGTAACAGATAACTGCTGAAGAAT 145
Db 1037 HTYWTTCWYMYWYCYCYCHHHWHMWMAYAWMMWHYAMWHHHHHMMCMCHWYTMH 978
QY 146 ACCGTTCGGCAGGCCCAAAACCCGTACTGCAACTTTTATTGTGAACCTCCCATTTATGAGA 205
Db 977 MCCVMYCCWMMWCHHMTATTWHYTYHYMMWCHTMTTWYHYMTMWHHHHAAM 918
QY 206 AAATCCCTTTTCGTCCTTTCTGTGTAATCGCTCCCTACTT 244
Db 917 AAHMHAATHWYHMAHAHTTYTYYTYYYYYYYYY 879

RESULT 14
EX456623/c
LOCUS
DEFINITION EX456623 Homo sapiens THYMUS Homo sapiens cdna clone CS0CAP003VC21
5-PRIME, mRNA sequence.
ACCESSION EX456623
VERSION EX456623
KEYWORDS EX456623.1 GI:31028793
SOURCE
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1173)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP003AB11QP1.
Location/Qualifiers

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 40.3413 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-92

Perfect score: 259

Sequence: 1 AATTGTTATATCAAGTAGG.....TACTTACTGCCAGCGAAATT 259

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.6	13.7	3499	3	US-08-857-076-43
2	33	12.7	2229	4	US-09-454-495-1
3	31	12.0	1553	4	US-09-461-325-74
4	30.4	11.7	148567	4	US-09-801-876B-3
5	29.6	11.4	5183	1	US-08-459-568-3
6	29.6	11.4	5183	2	US-08-399-411-3
7	29.6	11.4	5868	3	US-08-516-859A-3
8	29.6	11.4	5868	4	US-09-586-472-3
9	29.6	11.4	5868	4	US-09-528-706-3
10	29.4	11.4	5532	4	US-08-851-567B-52
11	29.4	11.4	6836	4	US-08-976-259-73
12	29.4	11.4	7515	4	US-08-851-567B-11
13	29.2	11.3	750	4	US-08-328-352-1687
14	29	11.2	2102	4	US-09-620-312D-487
15	29	11.2	10607	1	US-08-078-090-3
16	28.8	11.1	369	4	US-08-712-016-81
17	28.8	11.1	960	3	US-08-975-628-1
18	28.8	11.1	960	4	US-08-427-774-1
19	28.8	11.1	1113	4	US-09-134-001C-1556
20	28.8	11.1	3001	4	US-09-539-333D-131
21	28.8	11.1	3001	4	US-09-539-333D-137
22	28.8	11.1	1664976	4	US-08-916-421B-1
23	28.4	11.0	3308	3	US-08-714-918-68
24	28.4	11.0	3308	3	US-09-265-315-68
25	28.4	11.0	3308	3	US-09-265-315-68
26	28.4	11.0	3308	3	US-08-266-417-88
27	28.4	11.0	1230025	4	US-09-198-452A-1

C	28	28.2	10.9	767	4	US-09-634-238-177	Sequence 177, App
C	29	28.2	10.9	1001	4	US-09-671-317-427	Sequence 427, App
C	30	28.2	10.9	4008	4	US-09-134-001C-879	Sequence 879, App
	31	28.2	10.9	7218	1	US-08-232-463-14	Sequence 14, Appl
	32	28	10.8	2799	1	US-08-212-188-1	Sequence 1, Appl
	33	28	10.8	2799	3	US-08-970-725-1	Sequence 1, Appl
	34	28	10.8	2799	5	PCT-US95-02708-1	Sequence 1, Appl
	35	28	10.8	2967	4	US-09-614-480-1	Sequence 1, Appl
	36	28	10.8	3471	5	PCT-US93-00227-2	Sequence 2, Appl
C	37	28	10.8	112132	4	US-09-741-150-3	Sequence 3, Appl
C	38	27.8	10.7	827	5	PCT-US95-06406A-20	Sequence 20, Appl
C	39	27.8	10.7	1551	4	US-09-465-558-49	Sequence 49, Appl
C	40	27.6	10.7	800	4	US-09-222-938A-29	Sequence 29, Appl
	41	27.6	10.7	1177	3	US-08-861-774E-23	Sequence 23, Appl
	42	27.4	10.6	2822	2	US-08-678-405-1	Sequence 1, Appl
	43	27.4	10.6	2822	4	US-08-842-799-1	Sequence 1, Appl
	44	27.4	10.6	2822	5	PCT-US96-11458-1	Sequence 1, Appl
	45	27.2	10.5	559	4	US-09-340-323A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-08-857-076-43

Sequence 43, Application US/08857076C

Patent No. 6225120

GENERAL INFORMATION:

APPLICANT: Ruvkun, Gary

APPLICANT: Kimura, Koutarou

APPLICANT: Patterson, Garth

APPLICANT: Ogg, Scott

APPLICANT: Paradis, Suzanne

APPLICANT: Tissenbaum, Heidi

APPLICANT: Kowsek, Allison

TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR FILE REFERENCE: 00786/351001

CURRENT APPLICATION NUMBER: US/08/857,076C

CURRENT FILING DATE: 1997-05-15

NUMBER OF SEQ ID NOS: 114

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 43

LENGTH: 3499

TYPE: DNA

ORGANISM: Caenorhabditis elegans

US-08-857-076-43

Query Match 13.7%; Score 35.6; DB 3; Length 3499;

Best Local Similarity 52.7%; Pred. No. 0.069;

Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 27 ATTTATGCTGACCTACAAACCAAAACCACTACCACCTTATCACTCCACAACC 86

DB 2291 ATTTTCTTTTCTTCCAGATAGTAAATCATTCAGCTGAAATATTTGAAAAAAA 2350

QY 87 CTCTTCAGACACCTCGTTTTTTTGAACAAACATCTGTAACAGATAACTGCTGGAAGATA 146

DB 2351 CTTCAAATCAATTTTGTTCGAAAAAAAATTAATATATATATTTTCGAAATCTT 2410

QY 147 CGTTGCCGAGCCCAAAACCCGTAC 172

DB 2411 CGCTCTTCATCTTTCAATCCCTAC 2436

RESULT 2

US-09-454-495-1/c

Sequence 1, Application US/09454495

Patent No. 6576759

GENERAL INFORMATION:

APPLICANT: Reddy, Gurucharan

APPLICANT: Zeng, Hong

[illegible]

APPLICATION NUMBER: US 08/399,411


```

; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5158
US-08-459-568-3

Query Match 11.4%; Score 29.6; DB 1; Length 5183;
Best Local Similarity 57.6%; Pred. No. 7;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 CTCGTTTTTGAAGAACATCTGTAACAGATACTCTGAAGATAACCGTTCCGAGCC 159
Db 5002 CTCCTTTGTTCTCACTCAAGTCAGCAGCTGCCAGCTCGCGGTGACTGGCC 4943

QY 160 CCAAAACCGTACTGCAACTTTTATTGTGAAC 191
Db 4942 CCCCACTCCGCTCTCTAGATTATTTATTGAAC 4911

RESULT 6
US-08-399-411-3/c
; Sequence 3, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5158

```

```

US-08-399-411-3

Query Match 11.4%; Score 29.6; DB 2; Length 5183;
Best Local Similarity 57.6%; Pred. No. 7;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 CTCGTTTTTGAAGAACATCTGTAACAGATACTCTGAAGATAACCGTTCCGAGCC 159
Db 5002 CTCCTTTGTTCTCACTCAAGTCAGCAGCTGCCAGCTCGCGGTGACTGGCC 4943

QY 160 CCAAAACCGTACTGCAACTTTTATTGTGAAC 191
Db 4942 CCCCACTCCGCTCTCTAGATTATTTATTGAAC 4911

RESULT 7
US-08-516-859A-3/c
; Sequence 3, Application US/08516859A
; Patent No. 6069231
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516,859A
; FILING DATE: 18-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5868 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..5278
US-08-516-859A-3

Query Match 11.4%; Score 29.6; DB 3; Length 5868;
Best Local Similarity 57.6%; Pred. No. 7.4;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 CTCGTTTTTGAAGAACATCTGTAACAGATACTCTGAAGATAACCGTTCCGAGCC 159
Db 5122 CTCCTTTGTTCTCACTCAAGTCAGCAGCTGCCAGCTCGCGGTGACTGGCC 5063

QY 160 CCAAAACCGTACTGCAACTTTTATTGTGAAC 191

```

```
Db 5062 CCCACTCCGCTCTCTAGATTTTATTGAAC 5031
|||||
RESULT 8
US-09-586-472-3/c
; Sequence 3, Application US/09586472
; Patent No. 6323335
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/586,472
; FILING DATE: 01-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/528,706
; FILING DATE: 17-MAR-2000
; APPLICATION NUMBER: US 08/516,859
; FILING DATE: 18-AUG-1995
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 4130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5868 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..5278
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-586-472-3
Query Match 11.4%; Score 29.6; DB 4; Length 5868;
Best Local Similarity 57.6%; Pred. No. 7.4;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 100 CTCGTTTTTGA AAAACAATCTGTAACAGATACTGCTGAAGAAATACCGTTGCCGAGCC 159
Db 5122 CTCCTTTGTTCTCACTCAAGTCAGCAGCAGCTGCCAGCTGAAGGCTCCGGTGACTGGCC 5063
QY 160 CCAAAACCCGCTACTGCAACTTTTATTGTGAAC 191
Db 5062 CCCACTCCGCTCTCTAGATTTTATTGAAC 5031
|||||
RESULT 10
US-08-851-567B-52/c
; Sequence 52, Application US/08851567B
; Patent No. 6528484
; GENERAL INFORMATION:
; APPLICANT: Ensign, Jerald C
; APPLICANT: Bowen, David J
; APPLICANT: Petell, James
; APPLICANT: Fatig, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: ffrench-Constant, Richard
; APPLICANT: Rocheleau, Thomas A.
; APPLICANT: Blackburn, Michael B.
; APPLICANT: Hey, Timothy D.
```

```
; APPLICANT: Merlo, Donald J.
; APPLICANT: Orr, Gregory L.
; APPLICANT: Roberts, Jean L.
; APPLICANT: Strickland, James A.
; APPLICANT: Guo, Lining
; APPLICANT: Ciche, Todd A.
; APPLICANT: Sukhapinda, Kitisri
; TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,567B
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/063,615
; FILING DATE: 18-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/395,497
; FILING DATE: 28-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/007,255
; FILING DATE: 06-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,423
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/705,484
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seav, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.93804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-851-567B-52

Query Match 11.4%; Score 29.4; DB 4; Length 5532;
Best Local Similarity 58.6%; Pred. No. 8.4;
Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 124 AAACAGATACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTTTTA 183
DB 2828 ACATAGTTTCTGTATATAGACAGTTCAGAGACACCAGCCCAAGTACTGTACGTTTA 2769

QY 184 TTGTGAATTCCTTATGAGAAATC 210
DB 2768 TTGTAACTTCCAGTCAGTGAAGAAC 2742

RESULT 11
US-08-976-259-73
; Sequence 73, Application US/08976259
; Patent No. 6316609
```

```
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Choi, Gil H.
; APPLICANT: Welch, Rodney A.
; TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
; Patent No. 6316609
; NUMBER OF SEQUENCES: 142
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,259
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6836 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-976-259-73

Query Match 11.4%; Score 29.4; DB 4; Length 6836;
Best Local Similarity 56.8%; Pred. No. 9.1;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 126 ACAGATACTGCTGAAGATACCGTTGCCGAGCCCAAAACCGTACTGCAACTTTTAT 185
DB 292 ACTGATTGCAGTTCAGGATACCGTTTCCCAACACACACCGTGTTCATCATCTTTTAA 351

QY 186 GTGAATTCCTTATGAGAAATCCCTTTTCGTC 220
DB 352 TGTGAGGCACAGTACGAAACAAACAAAGTTTCATC 386

RESULT 12
US-08-851-567B-11/c
; Sequence 11, Application US/08851567B
; Patent No. 6526484
; GENERAL INFORMATION:
; APPLICANT: Ensign, Gerald C
; APPLICANT: Bowen, David J
; APPLICANT: Petell, James
; APPLICANT: Fatis, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: French-Constant, Richard
; APPLICANT: Rocheleau, Thomas A.
; APPLICANT: Blackburn, Michael B.
; APPLICANT: Hey, Timothy D.
; APPLICANT: Merlo, Donald J.
; APPLICANT: Orr, Gregory L.
; APPLICANT: Roberts, Jean L.
; APPLICANT: Strickland, James A.
; APPLICANT: Guo, Lining
; APPLICANT: Ciche, Todd A.
```

APPLICANT: Sukhapinda, Kitisri
TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46268

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,567B
FILING DATE: 05-MAY-1997

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J

REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.93804

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 7515 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..7515

US-08-851-567B-11

Query Match 11.4%; Score 29.4; DB 4; Length 7515;
Best Local Similarity 58.6%; Pred. No. 9.5;
Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 124 AACAGATACTGCTGAAGATAACCGTTCGCCAGCCCCCAAAACCCGCTACTGCAACTTTTA 183
Db 3089 ACATAGTTTCTGATATATAGACCACTTCAGAGACACCAGCCCAAGTACTGTAACTTTA 3030
QY 184 TTGTGAACCTCCCATTAATGAGAAATC 210
Db 3029 TTGTAACTCCCACTGCTGAGTGAAGAAC 3003

RESULT 13

US-09-328-352-1687/c

Sequence 1687, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 1687

LENGTH: 750

TYPE: DNA

ORGANISM: Acinetobacter baumannii

US-09-328-352-1687

Query Match 11.3%; Score 29.2; DB 4; Length 750;
Best Local Similarity 59.8%; Pred. No. 4.3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 82 AAACCTCTTCAGACAACCTCGTTTTTTGAAAAACAATCTGTAACACAGATAACTGCTGAA 141

Db 350 AAACCTTGCCTTAACCAAGCCATTTCTCAAAATAATATTTAAGCCCATCAGTCTGCTAA 291

QY 142 GAATACCGTTGCCGAGCCCCAA 163

Db 290 GAAGCCCCAAACATAACACCCA 269

RESULT 14

US-09-620-312D-487/c

Sequence 487, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John, Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt FL_genes Version 1.0

SEQ ID NO 487

LENGTH: 2102

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (309)..(1283)

US-09-620-312D-487

Query Match 11.2%; Score 29; DB 4; Length 2102;

Best Local Similarity 63.8%; Pred. No. 7.6;

Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 158 CCCCCAACCCGCTACTGCAACTTTTATGTGAACCTCCCATTTATGAGAAAATCCCTTTC 217

Db 461 CTCCAAGATGCTTCTGCACTTTTATATACATTCACCTGATTATAATTTTCCTTC 402

Qy 218 GTCCTCTTT 226
Db 401 AGACTCTGT 393

RESULT 15

US-08-078-090-3/c
; Sequence 3, Application US/08078090
; Patent No. 5739407
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, SVEN
; APPLICANT: HERNELL, OLLE
; APPLICANT: LOENNERDAL, BO
; APPLICANT: HJALMARSSON, KARIN
; APPLICANT: HANSSON, LENNART
; APPLICANT: TOERNELL, JAN
; APPLICANT: STROEMQUIST, MATS
; TITLE OF INVENTION: HUMAN BETA-CASEIN PROCESS FOR PRODUCING
; TITLE OF INVENTION: IT AND USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/078,090
; FILING DATE: 19930618
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DK92/00236
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DK91/00233
; FILING DATE: 19-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: BERGSTROM2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-5197
; TELEFAX: (202)737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(4804..4854, 5720..5746, 6726..6746, 6845
; LOCATION: ..6886, 7991..8521, 9440..9445)
US-08-078-090-3

Query Match 11.2%; Score 29; DB 1; Length 10607;
Best Local Similarity 53.0%; Pred. No.15;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 6 TGTATATCACTAGTAGGCGATTATGCTTACCTACCAAAACCAAAACCACTACCAC 65
Db 6674 TATAAATCAGGAATAAAGGCAATTAATCTTTGATATATCATCAATCACTTTGTATAT 6615
Qy 66 CTTTAATCAACTCCACAAACCTCTTTCAGACAACCTCGTTTTTTTGAACAAATCTG 122

Db 6614 CATTAACCTTATGTATTATTAACCTGTTCTATATGAGCTGGTTATGTAAAGAAAGTCAG 6558
Search completed: November 15, 2003, 08:09:36
Job time : 44.3413 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 146.667 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-92

Perfect score: 259

Sequence: 1 AATTGTTATATCAAGTAGG.....TACTTACTGCCAGCAAT 259

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/1/pubpna/PCCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/1/pubpna/PCCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	259	100.0	259	US-09-928-457-92
2	41.2	15.9	13326	Sequence 92, Appl
3	38	14.7	6944	Sequence 1686, Ap
4	38	14.7	16509	Sequence 112, App
5	37.8	14.6	5552	Sequence 1293, Ap
6	37	14.3	3673778	Sequence 1232, Ap
7	36.6	14.1	8576	Sequence 1, Appli
8	36.2	14.0	8781	Sequence 2201, Ap
9	35.8	13.8	7758	Sequence 2163, Ap
10	35.6	13.7	3499	Sequence 1660, Ap
11	35.6	13.7	3499	Sequence 1076, Ap
12	35.6	13.7	3499	Sequence 43, Appl
13	35.6	13.7	3499	Sequence 43, Appl
14	35.6	13.7	11996	Sequence 46, Appl
15	35.4	13.7	6534	Sequence 444, App
16	34.8	13.4	5929	Sequence 682, App

C 17	34.8	13.4	7491	12	US-10-311-455-1557	Sequence 1557, Ap
C 18	34.8	13.4	9293	12	US-10-240-453-33	Sequence 33, Appl
C 19	34.8	13.4	9293	14	US-10-239-676-25	Sequence 25, Appl
C 20	34.8	13.4	3673778	12	US-10-312-841-2	Sequence 2, Appli
C 21	34.4	13.3	6631	12	US-10-240-453-213	Sequence 213, App
C 22	34.4	13.3	9084	12	US-10-311-455-1578	Sequence 1578, Ap
C 23	34.4	13.3	11131	12	US-10-204-708-28	Sequence 28, Appl
C 24	34.4	13.3	11131	12	US-10-311-455-726	Sequence 726, App
C 25	34.4	13.3	13084	12	US-10-311-455-2151	Sequence 2151, Ap
C 26	34.2	13.2	7934	12	US-10-240-453-169	Sequence 169, App
C 27	34.2	13.2	7934	14	US-10-239-676-153	Sequence 153, App
C 28	34.2	13.2	13125	12	US-10-311-455-1200	Sequence 1200, Ap
C 29	34.2	13.2	13125	12	US-10-240-485-110	Sequence 110, App
C 30	34.2	13.2	14353	12	US-10-311-455-2407	Sequence 2407, Ap
C 31	34	13.1	6396	12	US-10-311-455-2220	Sequence 2220, Ap
C 32	34	13.1	17967	12	US-10-311-455-988	Sequence 988, App
C 33	33.8	13.1	5678	12	US-10-311-455-1111	Sequence 1111, Ap
C 34	33.8	13.1	8301	12	US-10-311-455-25	Sequence 25, Appl
C 35	33.8	13.1	12733	12	US-10-311-455-63	Sequence 63, Appl
C 36	33.6	13.0	764	12	US-10-027-632-150062	Sequence 150062,
C 37	33.6	13.0	764	12	US-10-027-632-150063	Sequence 150063,
C 38	33.6	13.0	764	12	US-10-027-632-150064	Sequence 150064,
C 39	33.6	13.0	764	12	US-10-027-632-150065	Sequence 150065,
C 40	33.6	13.0	764	13	US-10-027-632-150062	Sequence 150062,
C 41	33.6	13.0	764	13	US-10-027-632-150063	Sequence 150063,
C 42	33.6	13.0	764	13	US-10-027-632-150064	Sequence 150064,
C 43	33.6	13.0	764	13	US-10-027-632-150065	Sequence 150065,
C 44	33.6	13.0	4993	12	US-10-311-455-2076	Sequence 2076, Ap
C 45	33.4	12.9	6801	12	US-10-204-708-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-09-928-457-92

Sequence 92, Application US/09928457

Patent No. US2002016403A1

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: DNA, specific proteins and peptides

TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method for obtaining them and their biological application.

TITLE OF INVENTION: for obtaining them and their biological application.

NUMBER OF SEQUENCES: 99

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (OEB)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/928,457

FILING DATE: 2001-08-14

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/214,759

FILING DATE: 199-12-10

INFORMATION FOR SEQ ID NO: 92:

SEQUENCE CHARACTERISTICS:

LENGTH: 259 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHEetical: NO

ANTI-SENSE: NO

US-09-928-457-92

Query Match 100.0%; Score 259; DB 10; Length 259;

Best Local Similarity 100.0%; Pred. No. 3.2e-70;

Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTGTTATATCAAGTAGGCGCATTTGCTGACCTACAAACCAAAACAACCT 60

Db 1 AATTGTTATATCAAGTAGGCGCATTTGCTGACCTACAAACCAAAACAACCT 60

```
QY 61 ACCACCTTAATCAACTCCACAACCTCTTTCAGACAACCTCGTTTTTTGAAAAACAATC 120
DB 61 ACCACCTTTAATCAACTCCACAACCTCTTTCAGACAACCTCGTTTTTTGAAAAACAATC 120
QY 121 TGTAAACAGATAACTGCTGAAGATAACCGTTGCGAGCCCAACCCGCTACTGCAACTT 180
DB 121 TGTAAACAGATAACTGCTGAAGATAACCGTTGCGAGCCCAACCCGCTACTGCAACTT 180
QY 181 TTATTGTGAACCTCCCAATTATGAAAAATCCCTTTTCGCTCTCTTCTGTTATTCGTCCT 240
DB 181 TTATTGTGAACCTCCCAATTATGAAAAATCCCTTTTCGCTCTCTTCTGTTATTCGTCCT 240
QY 241 ACTTACTGCCAGCGAAATTT 259
DB 241 ACTTACTGCCAGCGAAATTT 259

RESULT 2
US-10-311-455-1686/c
; Sequence 1686, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1686
; LENGTH: 13326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1686

Query Match 15.9%; Score 41.2; DB 12; Length 13326;
Best Local Similarity 54.7%; Pred. No. 0.066;
Matches 82; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 36 TGACCTACAAACCAACCAACCTACCACCTTAATCACTCCACCAACCCCTCTTCAGA 95
DB 8480 TTATCTATAAACAATAACATATACCTCCCATCTCCACAACCAACCAACCACTCA 8421
QY 96 CAACCTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTGCTGAAGATAACCGTTGCGG 155
DB 8420 ACATATCTCATCTCTACAAAAAACCCTCAAAATTTCTCTTTTACCAACTCCCAATTAC 8361
QY 156 AGCCCAACCCGCTACTGCAACTTTTATT 185
DB 8360 TACCACATAACATACCTTTAAAAATAAT 8331

RESULT 3
US-10-172-086-112/c
; Sequence 112, Application US/10172086
; Publication No. US20030113750A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; TITLE OF INVENTION: of prostate tumors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
```

```
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 112
; LENGTH: 6944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-172-086-112

Query Match 14.7%; Score 38; DB 14; Length 6944;
Best Local Similarity 54.2%; Pred. No. 0.48;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 44 AAAACCAAAACAACCTTACCACCTTAATCACTCCACAACCCCTTTCAGACAACCTCG 103
DB 4871 AAATCCAAAAAACAACGACCTCTATAAAAAAATAAAAAACCCGCGATAAAAACGCA 4812
QY 104 TTTTGTAAAAACAATCTGTAAACAGATAACTGCTGAAGATAACCGTTGCGGAGCCCAA 163
DB 4811 AAACCTCAAAAAACCGCTCCGAAAAAATACTAACGAATCTCAACCCCTCTCGCCCCCA 4752
QY 164 AACCGCTACTGCAACTTTTATT 185
DB 4751 AACTCCCACTCTTAAAAAATTT 4730

RESULT 4
US-10-311-455-1293/c
; Sequence 1293, Application US/10311455
; Publication No. US20030143608A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1293
; LENGTH: 16509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1293

Query Match 14.7%; Score 38; DB 12; Length 16509;
Best Local Similarity 53.1%; Pred. No. 0.72;
Matches 103; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 41 TACAAAACCAAAACAACCTTACCACCTTAATCACTCCAC-AAACCTCTTCAGACAAC 99
DB 11520 TAAATATCAAAACTAACTACCACCTCTAAAAATGCCCAAAACCCCTTTATAACATA 11461
QY 100 CTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTGCTGAAGATAACCGTTGCGGAGCC 159
DB 11460 CCTATCTAAACAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11401
QY 160 CAAAACCCGCTACTGCAACTTTTATTGTGAACCTTCCCATTTAGAAAAATCCCTTTTCGT 219
DB 11400 TAAATCCGTTTCAAACTACTTTACTAAATATTCGACCTAAAAAATAATATCCCTAAACT 11341
QY 220 CCTCTTTCTGTATT 233
```


Db 11340 TTTTITTTTTTAAAT 11327

RESULT 5

US-10-311-455-1232/c

; Sequence 1232, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1232

; LENGTH: 5552

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1232

Query Match 14.6%; Score 37.8; DB 12; Length 5552;

Best Local Similarity 50.3%; Pred. No. 0.5;

Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 48 CCAAAAACAACCTACCACCTTATCACTCCACAACCCCTTCTCAGACACCTCGTTT 107

Db 4507 CAAATAAACCCCAACATCCTCTTAAAAATCCACTTAACCTATTAAATTTACTCGTAT 4448

QY 108 TTGAAAAACAATCTGTAACAGATAACTGCTGAAGATAACGTTGCCGAGCCCAAAACC 167

Db 4447 TTTAAATCTCTTAAATAATAAACAACCTCCAAACCTTTCCATTCAACCACTTAAACCT 4388

QY 168 CGTACTGCAACTTTTATTGTAACCTCCATTATGAGAAAATCCCTTTTGTCTCTTTTC 227

Db 4387 CAACCCCAATCAATTTCTTACACTTTTACTCTTAAAAATCCAAATCCCAATATACTTA 4328

QY 228 TGTAT 232

Db 4327 TCTTT 4323

RESULT 6

US-10-312-841-1/c

; Sequence 1, Application US/10312841

; Publication No. US20030186277A1

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC

; FILE REFERENCE: E01/1208/WO

; CURRENT APPLICATION NUMBER: US/10/312,841

; CURRENT FILING DATE: 2002-12-30

; NUMBER OF SEQ ID NOS: 2

; SEQ ID NO 1

; LENGTH: 3673778

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; NAME/KEY: unsure

; LOCATION: (3294164)

US-10-312-841-1

Query Match 14.3%; Score 37; DB 12; Length 3673778;

Best Local Similarity 57.3%; Pred. No. 18;

Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 30 TATGCTGACCTACAAAACCAAAACCAACCTACCACCTTAAATCACTCCACAACCCCTC 89

Db 3137528 TATCTTTTACCACACAATAATAAAAAAATAAACTCAAAAAAATTTAAAAAATTAC 3137469

QY 90 TTCAGACAACCTCGTTTTTTTGAAAAACAATCTGTAACAGATAACTGCTGAAGAATA 146

Db 3137468 CCAAAACCAATCACTATTATAAAACAATTAATTAATAAACAATTAAGTATA 3137412

RESULT 7

US-10-311-455-2201/c

; Sequence 2201, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 2201

; LENGTH: 8576

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-2201

Query Match 14.1%; Score 36.6; DB 12; Length 8576;

Best Local Similarity 51.5%; Pred. No. 1.4;

Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 35 CTGACCTACAAACCAAAACCAACCTACCACCTTAAATCACTCCACAACCCCTCTTCAG 94

Db 554 CTAAATTAACAATCTCTTAACCTTTTATCAAAATCTTAACTTCTTACATTAAATTACA 495

QY 95 ACAACCTCGTTTTTTTGAAAAACAATCTGTAACAGATAAATCTGCTGAAGAATACCGTTGCC 154

Db 494 ACATACCTCTTTAACTTAAAAAAATTTAATAATCCACCTTCTTAAACCTACTTCTATC 435

QY 155 GAGCCCCAAAACCCGTACTGCACTTTTATTGTGAACCTCCCA 197

Db 434 GATTCAACAACCTCATCTCCATCCAAATTTTATTCCTTACTA 392

RESULT 8

US-10-311-455-2163/c

; Sequence 2163, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

```
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2163
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2163

Query Match      14.1%; Score 36.4; DB 12; Length 2200;
Best Local Similarity 50.6%; Pred. No. 0.89;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 50 AAAAAACCACTGACACCTTAATCACTCCACAAACCCCTTTCAGACAACTCGTTTTT 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1468 AAAAAAATCAAAAAACACTAAAAATCTATAAAATTCGACAAACACAAACCGAAAC 1409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 110 GAAAAACCACTGTAACACAGATACTGCTGAAGATACTGCTGCGAGCCCAAAACCCG 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1408 TAAACGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 170 TACTGCAACTTTTATGTAACCTTCCCAATATGAGAAATCCCTTTTCGTCCTC 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1348 AACACCAACTTAAACCGTAAACTACGTAATACAAAAAACCCCGTTTCTCTCTC 1295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-311-455-1660/c
; Sequence 1660, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1660
; LENGTH: 8781
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1660

Query Match      14.0%; Score 36.2; DB 12; Length 8781;
Best Local Similarity 56.2%; Pred. No. 1.9;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 26 CATTTATGCTGACCTACAAACCAAAACCACTACACCCCTTAATCAACTCCACAAAC 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 708 CATCACTTCCAACTTAATATACCTCTCTCAATCTCCCTCGATCTCTCTTAC 649
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 86 CTTCTTCAGACAACTCGTTTTTGGAAAAACATCTGTAACAGATACTGCTGAAGAT 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 648 TCTTAAATCCCACTTTCTCTTTTAAAAAATAAACTTAAATATATAATACTAAAAAT 589
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 146 A 146

US-09-928-457-92.rnpb
Db 588 A 588

RESULT 10
US-10-311-455-1076/c
; Sequence 1076, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1076
; LENGTH: 7758
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1076

Query Match      13.8%; Score 35.8; DB 12; Length 7758;
Best Local Similarity 49.2%; Pred. No. 2.4;
Matches 94; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 41 TACAAACCAAAACCACTACCACTTAACTCCCAAAACCCCTTTCAGACAAAC 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3819 TACTAAAAAATAATACCTTAACACAAAAATACAAAAAATAATCATCTATATCC 3760
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 101 TCGTTTTTGAACAACATCTGTAACAGATACTGCTGAAGATACTGCTGCGAGCCC 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3759 CATCTTTTCTTATAACCATATTTTAAATTTATATAAAAAAATAATATAAAAAA 3700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 161 CAAAACCGCTACTGCACTTTTATGTGAACCTTCCCTTATGAGAAATCCCTTTTCGTC 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3699 AAAAAAATAACATTCACCTTCTTCTTAACTCTTAAATTTTAAACCAATCATTTA 3640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 221 CTCCTTCTGTA 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3639 CTTCTTTTATA 3629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-09-205-658-43
; Sequence 43, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205.658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
```

```
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-205-658-43

Query Match
Best Local Similarity 13.7%; Score 35.6; DB 9; Length 3499;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 27 ATTATGCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 86
DB 2291 ATTTTTCCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 2350
QY 87 CTCCTCAGACAACCTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTCTGAAGAATA 146
DB 2351 CTTCAATACAAATTTGTTTCGAAAAAAAATTTAAATATATATTTTCAGAAATCTT 2410
QY 147 CCGTTGCGGAGCCCAAAACCCGTAC 172
DB 2411 CCGTCTTCATCTTTTCAAAATCCCTAC 2436

RESULT 12
US-09-844-353A-43
; Sequence 43, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Koweek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351005
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-844-353A-43

Query Match
Best Local Similarity 13.7%; Score 35.6; DB 9; Length 3499;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 27 ATTATGCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 86
DB 2291 ATTTTTCCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 2350
QY 87 CTCCTCAGACAACCTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTCTGAAGAATA 146
DB 2351 CTTCAATACAAATTTGTTTCGAAAAAAAATTTAAATATATATTTTCAGAAATCTT 2410
QY 147 CCGTTGCGGAGCCCAAAACCCGTAC 172
DB 2411 CCGTCTTCATCTTTTCAAAATCCCTAC 2436

RESULT 13
US-09-963-693-43
; Sequence 43, Application US/09963693
; Patent No. US20030181364A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
```

```
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/963,693
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/205,658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857,076
; PRIOR FILING DATE: 1997-05-15
; PRIOR APPLICATION NUMBER: 08/888,534
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: US98/10080
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-963-693-43

Query Match
Best Local Similarity 13.7%; Score 35.6; DB 12; Length 3499;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 27 ATTATGCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 86
DB 2291 ATTTTTCCTGACCTACAAACCAAAACCAACCTACCAACCTTAATCAACTCCACAACC 2350
QY 87 CTCCTCAGACAACCTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTCTGAAGAATA 146
DB 2351 CTTCAATACAAATTTGTTTCGAAAAAAAATTTAAATATATATTTTCAGAAATCTT 2410
QY 147 CCGTTGCGGAGCCCAAAACCCGTAC 172
DB 2411 CCGTCTTCATCTTTTCAAAATCCCTAC 2436

RESULT 14
US-10-240-485-46/c
; Sequence 46, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 46
; LENGTH: 11996
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-46

Query Match
Best Local Similarity 13.7%; Score 35.6; DB 12; Length 11996;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
```

```
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 40 CTACAAAACCAAAACCACTTACCACCCCTTAATCAACTCCACAAACCCCTTTGAGACAAC 99
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1331 CTACACTAAATAAATAAGCGCATCATCCCTATTCTACAAAAAATAAACAACAAAAA 1272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 100 CTCGTTTTTTGAAAAACAATCTGTAAACAGATAACTGCTGAAGAATACCGTTGCCGAGCC 159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1271 CCCTACTATCTAAATTTAAACTTAAATACCTTAACCTCCATCTATTCTTAACTACTCC 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 160 CCAAAACCCGTACTGCAA 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1211 CCACACACCCCAATCAA 1194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-10-311-455-444/c
; Sequence 444, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 444
; LENGTH: 6534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-444

Query Match 13.78; Score 35.4; DB 12; Length 6534;
Best Local Similarity 50.94; Pred. No. 3;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
Qy 32 TGCCTGACCTACAAAACCAAAACCAACCTTACCACCCCTTAATCAACTCCACAAACCCCTCTT 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 TTCGCGACCTAAATAATCGTAAACCGAACCCTTACCTAAATCCGAACCTCTT 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 92 CAGACAACTCGTTTGTGAAAAACAAATCTGTAAACAGATAACTGCTGAAGAATACCGTT 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 AAAAAAACCGCCCTTTTAAATCTCTAATCTATTCAAAAAATAAACAACAAAAATAAATC 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 152 GCGAGCCCAAAACCGCTACTCACTTTTATGTGACTTCCC 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 TCCAACTATAAATCGGCTCATCTTAATATACTTATACTTCCC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: November 15, 2003, 08:32:35
Job time : 153.767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 1053.55 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-96

Perfect score: 238

Sequence: 1 AATTGATACGTTGGA...ATGAAACTGGAGCAATT 238

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	238	100.0	238	1	AF169474	AF169474 Neisseria
2	238	100.0	238	6	A68925	A68925 Sequence 96
3	238	100.0	238	6	BD063025	BD063025 DNA and s
4	238	100.0	349061	1	NMA223491	AL162753 Neisseria
C 5	233.2	98.0	10853	1	AE002528	AE002528 Neisseria
C 6	233.2	98.0	349980	6	AX044034	AX044034 Sequence
C 7	206.8	86.9	1155	6	AX024070	AX024070 Sequence
9	43.4	18.2	139189	10	AL935138	AL935138 Mouse DNA
10	43.4	18.2	195071	2	AC111049	AC111049 Mus muscu
11	43.4	18.2	318620	2	AC110910	AC110910 Mus muscu
12	43	18.1	11670	6	AX281284	AX281284 Sequence
C 13	43	18.1	143715	9	AL589947	AL589947 Human DNA
14	43	18.1	161828	2	AC021418	AC021418 Homo sapi
15	43	18.1	204122	2	AL450123	AL450123 Homo sapi
16	42.6	17.9	1364	8	YSCCAJ1	D21164 Saccharomyc
C 17	42.6	17.9	43100	8	SCB9379	U18796 Saccharomyc
C 18	41.8	17.6	7218	6	I66494	I66494 Sequence 14
19	41	17.2	130054	8	AC133862	AC133862 Medicago
C 20	41	17.2	139606	2	AC124957	AC124957 Medicago
21	41	17.2	202080	2	AC113906	AC113906 Rattus no
22	40.8	17.1	136105	9	AC023508	AC023508 Homo sapi
23	40.4	17.0	246163	2	AC121623	AC121623 Rattus no
C 24	40.2	16.9	128580	2	AP005466	AP005466 Oryza sat
C 25	40.2	16.9	158456	2	AP004731	AP004731 Oryza sat
C 26	40	16.8	11141	1	U67493	U67493 Methanococc
27	40	16.8	20933	6	AX458607	AX458607 Sequence
28	40	16.8	110000	6	AR271569_10	Continuation (11 o
29	39.8	16.7	6754	6	AX252005	AX252005 Sequence
30	39.8	16.7	6754	6	AX348778	AX348778 Sequence
C 31	39.8	16.7	248936	2	AC135200	AC135200 Rattus no
C 32	39.8	16.7	318115	2	AC098550	AC098550 Rattus no
C 33	39.6	16.6	12592	2	AX251796	AX251796 Sequence
C 34	39.6	16.6	147193	2	AC135914	AC135914 Oryza sat
C 35	39.6	16.6	250029	3	AE014820	AE014820 Plasmodiu
C 36	39.4	16.6	12901	1	AE011617	AE011617 Leptospir
C 37	39.4	16.6	157762	9	AP001893	AP001893 Homo sapi
38	39.4	16.6	161610	2	AC011728	AC011728 Homo sapi
C 39	39.4	16.6	168995	2	AP002359	AP002359 Homo sapi
40	39.4	16.6	180818	9	AP000821	AP000821 Homo sapi
41	39.2	16.5	6077	6	AX347120	AX347120 Sequence
C 42	39.2	16.5	110000	3	AC116984_0	AC116984 Dictyoste
C 43	39.2	16.5	330050	3	PFA929355	AL929355 Plasmodiu
44	39	16.4	145944	10	AC117259	AC117259 Mus muscu
45	39	16.4	169121	2	AC113475	AC113475 Mus muscu

ALIGNMENTS

RESULT 1
AF169474
LOCUS AF169474 238 bp DNA linear BCT 09-AUG-2000
DEFINITION Neisseria meningitidis strain Z2491 clone Em087 unknown sequence.
ACCESSION AF169474
VERSION AF169474.1 GI:9754682
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
Perrin A., Nassif X. and Tinsley C.R.
AUTHORS Identification of regions of the chromosome of Neisseria
TITLE meningitidis and Neisseria gonorrhoeae which are specific to

Pred. No. is the number of results predicted by chance to have a

```

pathogenic Neisseriae
Unpublished
REFERENCE 2 (bases 1 to 238)
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugirard, Paris 75015, France
FEATURES
    source
        Location/Qualifiers
            1..238
                /organism="Neisseria meningitidis"
                /mol_type="genomic DNA"
                /strain="22491"
                /db_xref="taxon:487"
                /clone="Em087"
BASE COUNT 85 a 18 c 64 g 71 t
ORIGIN
Query Match 100.0%; Score 238; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.6e-41;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Db 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Qy 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Qy 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Db 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Qy 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
Db 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
RESULT 2
LOCUS A68925 238 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 96 from Patent WO9802547.
ACCESSION A68925
VERSION A68925.1 GI:4759844
KEYWORDS
SOURCE unidentified
ORGANISM
    source
        Location/Qualifiers
            1..238
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
BASE COUNT 85 a 18 c 64 g 71 t
ORIGIN
Query Match 100.0%; Score 238; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.6e-41;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Db 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Qy 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Qy 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Db 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Qy 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
Db 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238

```

```

Qy 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Qy 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Db 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Qy 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
Db 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
RESULT 3
LOCUS BD063025 238 bp DNA linear PAT 27-AUG-2002
DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications.
ACCESSION BD063025
VERSION BD063025.1 GI:22608628
KEYWORDS JP 2001504684-A/87.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 238)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis species bacteria, method for obtaining them and their biological applications
JOURNAL Patent: JP 2001504684-A 87 10-APR-2001; INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN, SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/87 PD 10-APR-2001 PF 11-JUL-1997 JP 19980505685 PR 12-JUL-1996 FR 96/08768 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI CARLA VINALS, PI PETRA MERKER PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC Strandedness: Single; CC Topology: Linear; FH Key Location/Qualifiers
FEATURES
    source
        Location/Qualifiers
            1..238
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
BASE COUNT 85 a 18 c 64 g 71 t
ORIGIN
Query Match 100.0%; Score 238; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.6e-41;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Db 1 AATTGGATACGTTGGAAAGGGATATTTGATTGGGAATCGGATGAAGATAACGCTAGAT 60
Qy 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Qy 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Db 121 CTACTGGTTAACTTCAATCAGGGAACATCAAGAAAGTTAGGAAGTTGATTATCAGTTG 180
Qy 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
Db 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238

```

```

RESULT 4
NMA222491
LOCUS
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
ACCESSION NMA222491
VERSION AL162753.2
KEYWORDS AL162753.2 GI:7379120
SOURCE
ORGANISM
REFERENCE
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skellton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
PubMed 10761919
REFERENCE
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source
1. 349061
/organism="Neisseria meningitidis Z2491"
/mol_type="Genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/notes="serogroup: A"
complement(24..206)
repeat_unit
310..39"
/label="ATR"
209..212
/genes="NMA0368"
220..681
/genes="NMA0368"
220..681
/genes="NMA0368"
/notes="NMA0368, probable integral membrane protein, len:
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83669.1"
/db_xref="GI:7379121"
/db_xref="SPTREMBL:Q9JWH7"
/translation="MQEQNRKSPPIVMLLVSVLWLTASLSNVAFYLGNGHSGEGLTV
LILGSI PASIDIRYCAVYANYVWLAIVLLALRKVPVHFAAFWGLALVAFSVKAVYV
DEAGTSDIVRYGAGFYLWYAAFAVASIGTFAGNKKERKAASADGTGNDV"
671..1492
/genes="hemK"
671..1492
/genes="hemK"
/notes="NMA0369, hemK, HemK protein, len: 273 aa; similar
to e.g. HEMK_ECOLI_P37186 HEMK protein (277 aa), fasta
scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
/codon_start=1
/transl_table=11
/
product="very hypothetical protein NMA0372"
/
product="HemK protein"
/protein_id="CAB83670.1"
/db_xref="GI:7379122"
/db_xref="SPTREMBL:Q9JWH6"
/translation="MTEDKWLGLSKLPKNEARMLLOYVSEYTRVOLLTRGGEMPEDEV
RORAKLRRLNGEPVAYILGAREFYGRRTFVNFVSLIPRETHLVEAVLARLPEN
GRVWDLGTSGAVAVVALERPDADFASDISPPALETARKVAADLGARVEFYGSWF
DTDMPSEKWDIVSNPYIENGDLRKSQGLRFPETALDTFDSGLSCIRTLAQGAP
DLRARGFLLEHGFQDGAARVGLAENGFSVETPLDLAGLDRLVTLGKMKHLK"
831..840
/genes="hemK"
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1162..1171
/genes="hemK"
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1187..1207
/genes="hemK"
/notes="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
1565..1568
/genes="NMA0370"
1578..2966
/genes="NMA0370"
1578..2966
/genes="NMA0370"
/notes="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
P44640 hypothetical protein HI0325 (450 aa), fasta scores;
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83671.1"
/db_xref="GI:7379123"
/db_xref="SPTREMBL:Q9JWH5"
/translation="NNVVVAVIVMLVLSLRVHVVLSTLTGAFVGGAVAGMPLQNI
A
DAAGVQAGIIPFNKLEGAKIALSYMLGAFAMATHSGLPQQAAGAVVRKLN
R
GMGPDVSRSGEGVWKLLSILVNMNSQNIPIHIAFIPMIVPPLLIVFNRLKIDR
RLIACVITFGLVTTMYFLPYGFAIFLNEILIGNHSAAPOLDVKNINVAAMAIPAL
GMLAGLLAFVHYRPRLYQSNNDAGNADANRPPQSAVRSAAVAIAVCFAIQL
MYEDSLVGLMGLGEAVMLGVINEKDNKDFEGGICKMAMVGFMTIAAQGFPAVWNA
TCHIQPLVSSMAIFGNSKGMALMLVGLLVITMGICSSFSTLIIAIIYVPLCVGL
GFSPATVAIVGTAGALDAGSPASDSTLGTPTMGLNADGQHDHIDRSVLPITFIHNIP
LLIAGWIAAMVL"
2967..3191
/genes="slyX"
2967..3191
/genes="slyX"
/genes="slyX"
/notes="NMA0371, slyX, SLIX protein homolog, len: 74 aa;
similar to SLIX_HAEIN_P44759 SLIX protein homolog (73 aa),
fasta scores; E(): 0.0017, 33.8% identity in 74 aa
overlap, and SLIX_ECOLI_P30857 SLIX protein (72 aa), fasta
scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=11
/product="SLIX protein homolog"
/protein_id="CAB83672.1"
/db_xref="GI:7379124"
/db_xref="SPTREMBL:Q9JWH4"
/translation="MDAVQEPHRTLEIQLQALQEDVIAGLNAMVAELRQLDLQQA
QLRLLYQKMQDENPDQAQBPVSLRDEIPHY"
complement(3271..3423)
/genes="NMA0372"
complement(3271..3423)
/genes="NMA0372"
/notes="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=11

```

```

Db      326855 CTACTGGTTAACTCAATCAGGGAACATGAGAAAGTTAGGAAGTTGATTATTCAGTTG 326914
QY      181 ATAGAGTTAGATTCTTCTGGAACCTTTGTGAGGGATTCTATGAAAACTGGAAGCAATT 238
        |||||
        /translation="MFGRSLPLGSDGFFLAVVEYVPVGFICLQKGTAYFRSGGRPHRI
        KGILSD"
Db      326915 ATAGAGTTAGATTCTTCTGGAACCTTTGTGAGGGATTCTATGAAAACTGGAAGCAATT 326972

RESULT 5
AE002528      10853 bp      DNA      linear      BCT 25-MAY-2000
LOCUS      Neisseria meningitidis serogroup B strain MC58 section 170 of 206
DEFINITION
ACCESSION      AE002528
VERSION      AE002528.1 GI:7227034
KEYWORDS
SOURCE      Neisseria meningitidis MC58
ORGANISM      Neisseria meningitidis MC58
REFERENCE      1 (bases 1 to 10853)
AUTHORS      Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
              Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
              Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
              Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
              Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blait, E.,
              Ciftone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.,
              Qian, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
              Pizzu, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
              Rappuoli, R. and Venter, J.C.
              Complete genome sequence of Neisseria meningitidis serogroup B
              strain MC58
JOURNAL      Science 287 (5459), 1809-1815 (2000)
MEDLINE      20175755
PUBMED      10710307
REFERENCE      2 (bases 1 to 10853)
AUTHORS      Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
              Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
              Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
              Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
              Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
              Ciftone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.,
              Qian, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,
              Pizzu, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
              Rappuoli, R. and Venter, J.C.
              Direct Submission
TITLE      Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
JOURNAL      Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..10853
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/strain="MC58"
/db_xref="taxon:122586"
/notes="serogroup: B"
/codon_start=1
/complement(85..1827)
/gene="NM01780"
/complement(85..1827)
/gene="NM01780"
/notes="similar to PID:1621558 percent identity: 98.55;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hemolysin activation protein HecB, putative"
/protein_id="AAF42120.1"
/db_xref="GI:7227035"
/translation="MKFFPAPCLIVLIVLVLPLKTLAADENDABELIRSMQOQHIDAEI
LTDANVFPEOPLEKNVLSLEDPTCPVNVISLDKTVKFSFLPGLVLMKETAFTK
MCGSNLRLQKAAQOILVRGLYTSQAILIQPNMDSGLIKLRVAGSGIDIRYEEK
RDGKSAGSISAFNNKFLYRNKILNLRDVEQGLNLRPLPSYKTQIQLIPSEEGKS
DLQIKQONKAPFISIGIDAGGKTIGKYQGNVALSFDNPLGSDGTFVSYGRGLAHK
TDLTDTAGTETESGYSYVHYVPVKWLFSEFNHNRVHEATGFSVNVYDNGKQY
QSSLAARMRLNRNLKTSVGMKLWTKYKIDDAIEIVORRSAGWEALHRYAL
NRWOLDGKLSYKRTGMQSMPEENGDDILPQTSRMKTIITASLDAAPFILKQOF

Query Match      100.0%; Score 238; DB 1; Length 349061;
Best Local Similarity 100.0%; Pred. No. 6.7e-42;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AATTGGATACGTTGGAAGGATATTTGATTGGGAATGGGATGAAGTAACGCTAGAT 60
Db      326735 AATTGGATACGTTGGAAGGATATTTGATTGGGAATGGGATGAAGTAACGCTAGAT 326794
QY      61 GAGTTGGGAAAAGCTTAGACATATCGGTAAAGTAACCGCTATTGTTGAAAAT 120
Db      326795 GAGTTGGGAAAAGCTTAGACATATCGGTAAAGTAACCGCTATTGTTGAAAAT 326854
QY      121 CTACTGTTAACTCAATCAGGGAACATGAGAAAGTTAGGAAGTTGATTATTCAGTTG 180

```



```

FYATAIQANNKTLVAQDKLSIGSRVTVRGDGEQSLFGERGFYQWNTLTWYFHPNH
QFYLGADYGRVSGESAQYVSGKQLMGAVGFRGKHVGGMFAYDLFAGKFLHKKPGFQ
TNTVYGFNLNYS"
/translation="MSIHLDFKISPKTRQTYLYQKPLFKGAVRNLSASWKDINEI
YORAPLAPLHLRKKGAIVPEEYVESPDLLCKTRYRPIKSVIYEHMKNASLVNHN
INNEPDSHIAQVAFAGAGTIVSGVLAFGSDSEYKHNWDPRDYVATQLFGKRWQL
TAPDEPMYMOQXTDITDISEPHIDMDIILLAGGVLYIPIGWHMRPIPCGCTPHEA
VGTFFPENGYLEWLMKXFPPIESLRHSFSDWEQDRTINDTAAQIAMIADPVNVEA
FSDFLGKERTDIAHLEQFANPNATPLSDVRLKNANNDLTLEKGYLIDNGMKISV
DSLGGKHEHIGKNEFLULLKNLNVFNQGVVEVRKLIYQIULDFLEIL"
6633.7676
/gene="NM1787"
/codon_start=1
/transl_table=11
/product="N-acetyl-gamma-glutamyl-phosphate reductase"
/protein_id="AAF42126.1"
/db_xref="GI:7227041"
/translation="MSKIKVIGVATGYTGVELLRLAAHPDVEAAVTSRSEAGTA
VADYFSLRGVYGLAFQTPDEAGLECCDIVFATNGIAMKDAPELIQGVVRIDLSA
LPIRIDTWEHYGTHAAPDLVSGVAVGISELNREAVQAARLVANFGCYTCVLSLP
LVEPLRLKPGMPLIADCKSVGAGKNGVSLGCEAGNFAYGIAGHRHLPEI
ROTIAGLQDGIAREGFVFTPLAPMIRGMHATVYHLSDGSDPETVLRVYRDSPVDI
LPTGSAPSTRSVRGANLCEISITQQAQSDVWVLSVIDNLVKAAGQAVQNNNIMEGL
ETHGLDAIPLLP"
complement(8739..10781)
/gene="NM1788"
/codon_start=1
/transl_table=11
/product="ATP-dependent DNA helicase RecG"
/protein_id="AAF42127.1"
/db_xref="GI:7227042"
/translation="MMSETRKLKIDVSAKKDLKLNLTAWDLVHLPLRYEDETH
INPIDAPIGVPCQVEGEVHIEVTFKPKQLIVQIADGSGSVLFLRHFVASHQKQ
TAVKRIKRAVGIEKHGFYGDENIMHPIKDAEGGLAELTPVYVINGLNQTLRRII
OTALDVTPLHDTLPCLLKPLHPLHSLRLLHSPSPFTIHOISDGLTPAQWLKF
DELLAQQLSMRLAROKRIGTAAALGGDTLTQALQALFALTDAQKVEISEICRDM
AOTYPMHRLQDGVSGKTIYAALSAITAEISGAQVAVMAPEILAEQHFIFKQWLE
PLGIEVRLFGSLRKKDEAKLADGSKVIAVGHAFPSDGVAFHNLGITVBEQH
RGVQRALNKGREVQLKMSATPIERTLMSFPALDVSVIDELPGRTPIKTRL
VNNVRAEYGVLTGCRQAYVWCFLIEBETSLQJTAETLQALQTLALPELNIG
LVHGRMAAEAVNVARFSSGLNVLVATVIEGVDPVNAALYVIEHAERNGLAQLH
QURGVGAAESVCLVFAEPLGELAKARKVIEHTDGFIEARQDNLNIRGGBFLG
ARQSGVPMRLFAKLEEDLHLLEQARETAPMLIEQNPEIVEAHLARWLSREGYLVG"
BASE COUNT 2669 a 2735 c 2552 g 2897 t
ORIGIN
Query Match 98.0%; Score 233.2; DB 1; Length 10853;
Best Local Similarity 98.7%; Pred. No. 1.4e-40;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATTTGGATACGTTGGAAAAGGATATTGATTGGATGGCATGGAATAGCGTAGAT 60
Db 5155 AATTTGGATACGTTGGAAAAGGATATTGATTGGATGGCATGGAATAGCGTAGAT 5096
QY 61 GAGTTGGGAAAAAGTGTAGAACATATCGGTAGAAATGAACCGCTATTGTTGAAAAA 120
Db 5095 GAGTTGGGAAAAAGTGTAGAACATATCGGTAGAAATGAACCGCTATTGTTGAAAAA 5036
QY 121 CTACTGGTTAACTTCAATCAGGAAACATGAGAGTAGTAGAAGTTGATTTATCAGTGT 180
Db 5035 CTACTGGTTAACTTCAATCAGGAAACATGAGAGTAGTAGAAGTTGATTTATCAGTGT 4976
QY 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTCATGAAAACTGGAAGCAATT 238
Db 4975 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTCATGAAAACTGGAAGCAATT 4918

```

```

FYATAIQANNKTLVAQDKLSIGSRVTVRGDGEQSLFGERGFYQWNTLTWYFHPNH
QFYLGADYGRVSGESAQYVSGKQLMGAVGFRGKHVGGMFAYDLFAGKFLHKKPGFQ
TNTVYGFNLNYS"
/translation="MSIHLDFKISPKTRQTYLYQKPLFKGAVRNLSASWKDINEI
YORAPLAPLHLRKKGAIVPEEYVESPDLLCKTRYRPIKSVIYEHMKNASLVNHN
INNEPDSHIAQVAFAGAGTIVSGVLAFGSDSEYKHNWDPRDYVATQLFGKRWQL
TAPDEPMYMOQXTDITDISEPHIDMDIILLAGGVLYIPIGWHMRPIPCGCTPHEA
VGTFFPENGYLEWLMKXFPPIESLRHSFSDWEQDRTINDTAAQIAMIADPVNVEA
FSDFLGKERTDIAHLEQFANPNATPLSDVRLKNANNDLTLEKGYLIDNGMKISV
DSLGGKHEHIGKNEFLULLKNLNVFNQGVVEVRKLIYQIULDFLEIL"
6633.7676
/gene="NM1787"
/codon_start=1
/transl_table=11
/product="N-acetyl-gamma-glutamyl-phosphate reductase"
/protein_id="AAF42126.1"
/db_xref="GI:7227041"
/translation="MSKIKVIGVATGYTGVELLRLAAHPDVEAAVTSRSEAGTA
VADYFSLRGVYGLAFQTPDEAGLECCDIVFATNGIAMKDAPELIQGVVRIDLSA
LPIRIDTWEHYGTHAAPDLVSGVAVGISELNREAVQAARLVANFGCYTCVLSLP
LVEPLRLKPGMPLIADCKSVGAGKNGVSLGCEAGNFAYGIAGHRHLPEI
ROTIAGLQDGIAREGFVFTPLAPMIRGMHATVYHLSDGSDPETVLRVYRDSPVDI
LPTGSAPSTRSVRGANLCEISITQQAQSDVWVLSVIDNLVKAAGQAVQNNNIMEGL
ETHGLDAIPLLP"
complement(8739..10781)
/gene="NM1788"
/codon_start=1
/transl_table=11
/product="ATP-dependent DNA helicase RecG"
/protein_id="AAF42127.1"
/db_xref="GI:7227042"
/translation="MMSETRKLKIDVSAKKDLKLNLTAWDLVHLPLRYEDETH
INPIDAPIGVPCQVEGEVHIEVTFKPKQLIVQIADGSGSVLFLRHFVASHQKQ
TAVKRIKRAVGIEKHGFYGDENIMHPIKDAEGGLAELTPVYVINGLNQTLRRII
OTALDVTPLHDTLPCLLKPLHPLHSLRLLHSPSPFTIHOISDGLTPAQWLKF
DELLAQQLSMRLAROKRIGTAAALGGDTLTQALQALFALTDAQKVEISEICRDM
AOTYPMHRLQDGVSGKTIYAALSAITAEISGAQVAVMAPEILAEQHFIFKQWLE
PLGIEVRLFGSLRKKDEAKLADGSKVIAVGHAFPSDGVAFHNLGITVBEQH
RGVQRALNKGREVQLKMSATPIERTLMSFPALDVSVIDELPGRTPIKTRL
VNNVRAEYGVLTGCRQAYVWCFLIEBETSLQJTAETLQALQTLALPELNIG
LVHGRMAAEAVNVARFSSGLNVLVATVIEGVDPVNAALYVIEHAERNGLAQLH
QURGVGAAESVCLVFAEPLGELAKARKVIEHTDGFIEARQDNLNIRGGBFLG
ARQSGVPMRLFAKLEEDLHLLEQARETAPMLIEQNPEIVEAHLARWLSREGYLVG"
BASE COUNT 2669 a 2735 c 2552 g 2897 t
ORIGIN
Query Match 98.0%; Score 233.2; DB 1; Length 10853;
Best Local Similarity 98.7%; Pred. No. 1.4e-40;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATTTGGATACGTTGGAAAAGGATATTGATTGGATGGCATGGAATAGCGTAGAT 60
Db 5155 AATTTGGATACGTTGGAAAAGGATATTGATTGGATGGCATGGAATAGCGTAGAT 5096
QY 61 GAGTTGGGAAAAAGTGTAGAACATATCGGTAGAAATGAACCGCTATTGTTGAAAAA 120
Db 5095 GAGTTGGGAAAAAGTGTAGAACATATCGGTAGAAATGAACCGCTATTGTTGAAAAA 5036
QY 121 CTACTGGTTAACTTCAATCAGGAAACATGAGAGTAGTAGAAGTTGATTTATCAGTGT 180
Db 5035 CTACTGGTTAACTTCAATCAGGAAACATGAGAGTAGTAGAAGTTGATTTATCAGTGT 4976
QY 181 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTCATGAAAACTGGAAGCAATT 238
Db 4975 ATAGAGTTAGATTTTCTGGAACCTTTTGTGAGGATTCATGAAAACTGGAAGCAATT 4918

```

```

RESULT 6
AX044034/c
LOCUS      AX044034      349980 bp      DNA      linear      PAT 24-NOV-2000
DEFINITION Sequence 113 from Patent WO0066791.
ACCESSION  AX044034
VERSION     AX044034.1  GI:11342918
KEYWORDS
SOURCE
ORGANISM   Neisseria meningitidis
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE  1
AUTHORS    Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C.,
            Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
            Scarlato, V., Rapicotti, R., Frazer, C.M. and Grandi, G.
TITLE      Neisseria genomic sequences and methods of their use
JOURNAL    Patent: WO 0066791-A 113 09-NOV-2000;
            CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES   Location/Qualifiers
            source
            1..349980
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /db_xref="taxon:487"
            /note="sequence too long, cut in 8 pieces -seq 1: 1 to
            349980 349980 bases-seq 108: 300001 to 649980 349980
            bases-seq 109: 600001 to 949980 349980 bases-seq 110:
            900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
            349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
            113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
            2272325 172325 bases"
BASE COUNT 86473 a 35646 c 85908 g 81953 t
ORIGIN
Query Match      98.0%; Score 233.2; DB 6; Length 349980;
Best Local Similarity 98.7%; Pred. No. 7.2e-41;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATTTGGATACGTTGGAAAAGGGATATTGATTTGGGATCGGATGAGTAACCGTAGAT 60
|||||
Db 72678 AATTGGATACGTTGGAAAAGGGATATTGATTTGGGATCGGATGAGTAACCGTAGAT 72619

QY 61 GAGTTGGGAAAAAAGTTGTAGACATATCGTAAAGAAATCAACCGTTATTGTTGAAAAAT 120
|||||
Db 72618 GAGTTGGGAAAAAAGTTGTAGACATATCGTAAAGAAATCAACCGTTATTGTTGAAAAAT 72559

QY 121 CTACTGGTTAACTCAATCAGGAAAAACATGAAAGATTAGGAAGTTGATTTATCAGTTG 180
|||||
Db 72558 CTACTGGTTAACTCAATCAGGAAAAATATGAAAGATTAGGAAGTTGATTTATCAGTTG 72499

QY 181 ATAGAGTTAGATTTTCTGGAACCTTTCTGAGGGATTCTATGAAAAAAGTGAAGCAATT 238
|||||
Db 72498 ATAGAGTTAGATTTTCTGGAACCTTTCTGAGGGATTCTATGAAAAAAGTGAAGCAATT 72441

RESULT 7
AX024070
LOCUS      AX024070      1155 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION Sequence 13 from Patent FR2785293.
ACCESSION  AX024070
VERSION     AX024070.1  GI:10184382
KEYWORDS
SOURCE
ORGANISM   Neisseria meningitidis
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE  1
AUTHORS    Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
            Bouchardon, A. and Renaud, M.G.
JOURNAL    Patent: FR 2785293-A 13 05-MAY-2000;
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES   Location/Qualifiers
            source
            1..1155
            /organism="Neisseria meningitidis"
            /mol_type="genomic DNA"
            /db_xref="taxon:487"
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAC08934.1"
            /db_xref="GI:10184429"
            /translation="WGHLDPGIGSPKTFROTLYQKPKFKGAVRNLEAASCKYINEL
            YQRADPTAPLFLKRGKGAIVPKVEVSFDDLGKTRVRFIKSVIYEHKNGASLVYNH
            INNEPFSHDHARQVARFAGAHITVSGYLAQSDSYKHNWDRDVAIQLFGRKRWQL
            TAPDFPMPLYMQQTKDIDISPEIDMDIILEAGDVLYIPRGWHRHPIPGCETPFA
            VGTFFPNGYNYLEWLMKKFPTIESLRHSFSDWEQDRTRINDTAAQIAAMIAADPYNIEA
            FSEDFLQKERTDTAFHLEQFANPATPLSDVLRNLNANLDTLEKGYLIGNMKISV
            DELGKKVLEHIGKNEPLLKLLNLFNFGKHEEVKRLIYQLIELDFLEL"
BASE COUNT 335 a 269 c 248 g 303 t
ORIGIN
Query Match      86.9%; Score 206.8; DB 6; Length 1155;
Best Local Similarity 99.0%; Pred. No. 9.8e-35;
Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATTTGGATACGTTGGAAAAGGGATATTGATTTGGGATCGGATGAGTAACCGTAGAT 60
|||||
Db 946 AATTTGGATACGTTGGAAAAGGGATATTGATTTGGGATCGGATGAGTAACCGTAGAT 1005

QY 61 GAGTTGGGAAAAAAGTTGTAGACATATCGTAAAGAAATCAACCGTTATTGTTGAAAAAT 120
|||||
Db 1006 GAAATTTGGGAAAAAAGTTGTAGACATATCGTAAAGAAATCAACCGTTATTGTTGAAAAAT 1065

QY 121 CTACTGGTTAACTCAATCAGGAAAAACATGAAAGATTAGGAAGTTGATTTATCAGTTG 180
|||||
Db 1056 CTACTGGTTAACTCAATCAGGAAAAACATGAAAGATTAGGAAGTTGATTTATCAGTTG 1125

QY 181 ATAGAGTTAGATTTTCTGGAACCTTTCTGGAACCTTTTGTGA 210
|||||
Db 1126 ATAGAGTTAGATTTTCTGGAACCTTTTGTGA 1155

RESULT 8
AX024117
LOCUS      AX024117      1155 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION Sequence 60 from Patent FR2785293.
ACCESSION  AX024117
VERSION     AX024117.1  GI:10184428
KEYWORDS
SOURCE
ORGANISM   Neisseria gonorrhoeae
            Neisseria gonorrhoeae
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE  1
AUTHORS    Nassif, X., Tinsley, C., Aujame, L., Perrin, A., Rokbi, B.,
            Bouchardon, A. and Renaud, M.G.
JOURNAL    Patent: FR 2785293-A 60 05-MAY-2000;
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES   Location/Qualifiers
            source
            1..1155
            /organism="Neisseria gonorrhoeae"
            /mol_type="genomic DNA"
            /db_xref="taxon:485"
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAC08934.1"
            /db_xref="GI:10184429"
            /translation="WGHLDPGIGSPKTFROTLYQKPKFKGAVRNLEAASCKYINEL
            YQRADPTAPLFLKRGKGAIVPKVEVSFDDLGKTRVRFIKSVIYEHKNGASLVYNH
            INNEPFSHDHARQVARFAGAHITVSGYLAQSDSYKHNWDRDVAIQLFGRKRWQL
            TAPDFPMPLYMQQTKDIDISPEIDMDIILEAGDVLYIPRGWHRHPIPGCETPFA
            VGTFFPNGYNYLEWLMKKFPTIESLRHSFSDWEQDRTRINDTAAQIAAMIAADPYNIEA
            FSEDFLQKERTDTAFHLEQFANPATPLSDVLRNLNANLDTLEKGYLIGNMKISV
            DELGKKVLEHIGKNEPLLKLLNLFNFGKHEEVKRLIYQLIELDFLEL"

```

FSDFPLGKERTDTAFHLEQFANPNATPLSDVRLRLNANNLDLTKYLIQNGMKISV
DELGKVLHIGKNEPULLKNLVNFQAQHEEVKRLIYQILDFLEIL"

BASE COUNT
ORIGIN

337 a 265 c 246 g 307 t

Query Match 85.5%; Score 203.6; DB 6; Length 1155;
Best Local Similarity 98.1%; Pred. No. 4.8e-34; Indels 0; Gaps 0;
Matches 206; Conservative 0; Mismatches 4;
QY 1 AATTGGATAGCTGGAAAAGGGATATTGTTGGGAATGGATCAAGATAAGCGTAGAT 60
DB 946 AATTGGATAGCTGGAAAAGGGATATTGTTGGGAATGGATCAAGATAAGCGTAGAT 1005
QY 61 GAGTTGGGAAAAAGTGTAGACATATCGTTAGAAATGAACCGTTATTGTTGAAAAAT 120
DB 1006 GAGTTGGGAAAAAGTGTAGACATATCGTTAGAAATGAACCGTTATTGTTGAAAAAT 1065
QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATCAAGAAAGTTAGAAAGTTGATTATCAGTTG 180
DB 1056 CTACTGGTTAACTTCAATCAGGAAAAACATCAAGAAAGTTAGAAAGTTGATTATCAGTTG 1125
QY 181 ATAGAGTTAGATTTCTCGAACTTTGTGA 210
DB 1126 ATAGAGTTAGATTTCTCGAAATTTGTGA 1155

RESULT 9
AL935138
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism

Mouse DNA sequence from clone RP23-222H15 on chromosome 2, complete
sequence.
AL935138
GI:27374701
HTG
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (21-DEC-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Dec 26, 2002 this sequence version replaced gi:26985395.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the rare
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep RP23-222H15 is
from the RP23-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6

FEATURES
source

Location/Qualifiers
1..139189
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="2"
/clone="RP23-222H15"
/clone_lib="RP23-23"

BASE COUNT 42936 a 24714 c 25557 g 46082 t
ORIGIN

Query Match 18.2%; Score 43.4; DB 10; Length 139189;
Best Local Similarity 64.4%; Pred. No. 4.8;
Matches 65; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 118 ATCTACTGGTTAACTTCAATCAGGAAAAACATCAAGAAAGTTAGAAAGTTGATTATCAG 177
DB 89602 ATTCTATTTTAAATTTTATCTTTAAAGCCTTAAGCAATGCAAAATGTTTATAAG 89661
QY 178 TTGATAGAGTTAGATTTCTGGAACCTTTGTGAGGATTTCT 218
DB 89662 TTATATAGATAGAAATTAAGAGCTTTTGTAAATTTCT 89702

RESULT 10
AC111049

LOCUS
DEFINITION

Mus musculus clone RP23-421C14, WORKING DRAFT SEQUENCE, 6 ordered
pieces.
AC111049
GI:24431747
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS
SOURCE

Mus musculus (house mouse)
Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195071)

REFERENCE
AUTHORS

Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-421C14
Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 195071)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguski, M., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choquel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Govette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Huine, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R.,
Landers, T., Lechoczky, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 195071)
REFERENCE
AUTHORS

Barna.N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Rohr,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Thomann,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (31-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Oct 31, 2002 this sequence version replaced gi:21362181.

All repeats were identified using RepeatMasker:

Smit, A.F.P. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19878

Center clone name: 421.C.14

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 193450 bases at least Q40

Consensus quality: 193982 bases at least Q30

Consensus quality: 194232 bases at least Q20

Insert size: 183000; agarose-fp

Insert size: 194571; sum-of-contigs

Quality coverage: 9.5 in Q20 bases; agarose-fp

Quality coverage: 8.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 6 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 8394: contig of 8394 bp in length

* 8395 8494: gap of 100 bp

* 8495 9191: contig of 897 bp in length

* 9192 9291: gap of 100 bp

* 9292 121914: contig of 112623 bp in length

* 121915 122014: gap of 100 bp

* 122015 144276: contig of 22262 bp in length

* 144277 144376: gap of 100 bp

* 144377 185733: contig of 41357 bp in length

* 185734 185833: gap of 100 bp

* 185834 195071: contig of 9238 bp in length.

Location/Qualifiers

1. 195071

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="RP23-421C14"

/clone_lib="RPCI-23 Female Mouse BAC"

1. 8394

/note="assembly_fragment"

FEATURES

source

misc_feature

clone_end:SP6
vector_side:left
8495..19191
/note="assembly_fragment"
misc_feature
9292..121914
/note="assembly_fragment"
misc_feature
122015..144276
/note="assembly_fragment"
misc_feature
144377..185733
/note="assembly_fragment"
misc_feature
185834..195071
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 63730 a 34639 c 34804 g 61398 t 500 others
ORIGIN

Query Match 18.2%; Score 43.4; DB 2; Length 195071;
Best Local Similarity 64.4%; Pred No. 4.5;
Matches 65; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 118 AATCTACTGGTTACTTCAATCAGGAGAAACATGAAGAAGTTAGGAAGTTGATTATCAG 177
Db 48286 ATTCTATTTTAAATTTTATATCTTAAAGCCTAAAGCAATGACAAAATTTCTTTTATAAG 48345

QY 178 TTGATAGAGTTAGATTCTTCTGGAACCTTTTGTGAGGGATTCT 218
Db 48346 TTAATATAGATAGAAATTATGAAGCTTTTGCTAAAAATCT 48386

RESULT 11
AC110910

LOCUS AC110910 318620 bp DNA linear HTG 23-APR-2003
Mus musculus clone RP24-349E9, WORKING DRAFT SEQUENCE, 23 unordered pieces.

ACCESSION AC110910
VERSION AC110910.2 GI:30018095
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 318620)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-349E9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 318620)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (16-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 318620)

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouilleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boughalter,B., Camarata,J., Chang,J., Choepel,Y., Collimore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Haggopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,K., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,A., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schubback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 17, 2003 this sequence version replaced g1:19693473.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L21261

Center clone name: 349_E_9

* NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1220: contig of 1220 bp in length

* 1221 1320: gap of 100 bp

* 1321 2507: contig of 1187 bp in length

* 2508 2607: gap of 100 bp

* 2608 4770: contig of 2163 bp in length

* 4771 4870: gap of 100 bp

* 4871 5938: contig of 1068 bp in length

* 5939 6038: gap of 100 bp

* 6039 7850: contig of 1812 bp in length

* 7851 7950: gap of 100 bp

* 7951 10235: contig of 2285 bp in length

* 10236 10335: gap of 100 bp

* 10336 13293: contig of 2958 bp in length

* 13294 13393: gap of 100 bp

* 13394 16284: contig of 2891 bp in length

* 16285 16384: gap of 100 bp

* 16385 19117: contig of 2733 bp in length

* 19118 19217: gap of 100 bp

* 19218 23060: contig of 3843 bp in length

* 23061 23160: gap of 100 bp

* 23161 29199: contig of 6039 bp in length

* 29200 29299: gap of 100 bp

* 29300 36969: contig of 7670 bp in length

* 36970 37069: gap of 100 bp

* 37070 44337: contig of 7268 bp in length

* 44338 44437: gap of 100 bp

* 44438 54643: contig of 10206 bp in length

* 54644 54743: gap of 100 bp

* 54744 64252: contig of 9509 bp in length

* 64253 64352: gap of 100 bp

* 64353 76528: contig of 12176 bp in length

* 76529 76629: gap of 100 bp

* 76630 90709: contig of 13981 bp in length

* 90710 90709: gap of 100 bp

* 90711 111194: contig of 20485 bp in length

* 111195 111294: gap of 100 bp

* 111295 126721: contig of 15427 bp in length

* 126722 126821: gap of 100 bp

* 126822 147846: contig of 21025 bp in length

* 147847 147946: gap of 100 bp

* 147947 173784: contig of 25838 bp in length

* 173785 173884: gap of 100 bp

* 173885 217072: contig of 43188 bp in length

* 217073 217172: gap of 100 bp

* 217173 318620: contig of 101448 bp in length.

FEATURES

Location/Qualifiers

source 1..318620

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone_lib="RP24-349E9"

/clone_name="RP24-349E9"

/note="assembly_fragment"

1..1220

/note="assembly_fragment"

clone_end:Sp6

vector_side:left"

misc_feature 1321..2507

/note="assembly_fragment"

misc_feature 2608..4770

/note="assembly_fragment"

misc_feature 4871..5938

/note="assembly_fragment"

misc_feature 6039..7850

/note="assembly_fragment"

misc_feature 7951..10235

/note="assembly_fragment"

misc_feature 10336..13293

/note="assembly_fragment"

misc_feature 13394..16284

/note="assembly_fragment"

misc_feature 16385..19117

/note="assembly_fragment"

misc_feature 19218..23060

/note="assembly_fragment"

misc_feature 23161..29199

/note="assembly_fragment"

misc_feature 29300..36969

/note="assembly_fragment"

misc_feature 37070..44337

/note="assembly_fragment"

misc_feature 44438..54643

/note="assembly_fragment"

misc_feature 54744..64252

/note="assembly_fragment"

misc_feature 64353..76528

/note="assembly_fragment"

misc_feature 76629..90609

/note="assembly_fragment"

misc_feature 90710..111194

/note="assembly_fragment"

misc_feature 111295..126721

/note="assembly_fragment"

misc_feature 126822..147846

/note="assembly_fragment"

misc_feature 147947..173784

/note="assembly_fragment"

misc_feature 173885..217072

/note="assembly_fragment"

misc_feature 217173..318620

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 97164 a 60953 c 59588 g 98713 t 2202 others
ORIGIN

Query Match 19.2%; Score 43.4; DB 2; Length 318620;
Best Local Similarity 64.4%; Pred. No. 4.1;
Matches 65; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 118 AATCTACTGGTTAACTTCAATTCAGGGGAAACATGAGAGCTTAGGAAGTTAGTATTATATCAG 177
|||
Db 212850 ATTCTATTTTAAATTTTATCTTAAAGCCCTAAGCAATGACAAATTTGTTTTATAAG 212909
|||
Qy 178 TTGATAGAGTTAGATTTTCTGGAACTTTTGTGAGGGATTCT 218
|||
Db 212910 TTAATATAGATAGAAATATGAGCTTTTGTGTAATAATTTCT 212950
|||

RESULT 12
AX281284 AX281284 11670 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 26 from Patent WO0177164.
ACCESSION AX281284
VERSION AX281284.1 GI:16608539
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Olek.A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with apoptosis
JOURNAL Patent: WO 0177164-A 26 18-OCT-2001;
Epigenomics AG (DE)

FEATURES
Location/Qualifiers
source
1. 11670
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 3563 a 238 c 2443 g 5426 t
ORIGIN
Query Match 18.1%; Score 43; DB 6; Length 11670;
Best Local Similarity 49.3%; Pred. No. 9.4;
Matches 112; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 3 TTGGATACGTTGGAAGGCGATATTTCATTCGGATGGATGGAATGAAGTAAAGCTAGATCA 62
|||
Db 265 TTAGGTTATGATTGAAAATATATTTTATATATGATTTATATTAATAAATTTTGAAGTGGG 324
|||
Qy 63 GTTGGGAAAAAGTGTGAGACATATCGGTGAAGATGAACCTTATTCTTGAAAAATCT 122
|||
Db 325 GTAAGAAATAAAATGTTTTTATTATAGATTATTAATTTTGTGTTTGGATA 384
|||
Qy 123 ACTGGTTAACTCAATCAGGAAACATGAAGATTAGAGTTGATTTATCAGTTGAT 182
|||
Db 385 GAGGTTGATTTAGATCGTAATAATAGTAATTAATAATTAATTAATTAATTAATTAATTAAT 444
|||
Qy 183 AGAGTTAGATTTCTGGAACCTTTTGTGAGGATTCATGAAAACCTG 229
|||
Db 445 AGAGTTAGGTTGGTGTGTTTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 491
|||

RESULT 13
AL589947/c AL589947 143715 bp DNA linear PRI 27-APR-2001
LOCUS Human DNA sequence from clone RP11-127B16 on chromosome 6, complete
DEFINITION sequence.
ACCESSION AL589947 AC064822
VERSION AL589947.3 GI:13897193
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Almeida, J.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On or before May 15, 2001 this sequence version replaced
gi:7801478, gi:13751578.

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RP11-127B16 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-127B16. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-538A16 is at 143616 in this sequence. The true right end of clone RP11-596A13 is at 100 in this sequence.

FEATURES
source

Location/Qualifiers

1. 143715
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-127B16"
/clone_lib="RPCI-11.1"
954..1833
/note="L1PA7 repeat: matches 5256..6140 of consensus"
2919..3386
/note="L1R1D repeat: matches 1..505 of consensus"
3803..4812
/note="pTRS repeat: matches 742..1676 of consensus"
3817..4406
/note="CpG island"
/evidence=not_experimental
4813..5485
/note="L1R12 repeat: matches 1..671 of consensus"
5486..5500
/note="pTRS repeat: matches 1676..1690 of consensus"
5981..6327
/note="AluY repeat: matches 3..292 of consensus"
6608..6759
/note="76 copies 2 mer tt 62% conserved"
7904..8059
/note="L1NA5 repeat: matches 6137..6296 of consensus"
9414..12009
/note="L1PA6 repeat: matches 3557..6143 of consensus"
12010..12317
/note="AluYvA5 repeat: matches 1..308 of consensus"
12318..12812
/note="L1PA6 repeat: matches 3067..3557 of consensus"

```
repeat_region 12808..13670
/note="L1Pa6 repeat: matches 2194..3061 of consensus"
repeat_region 15757..15804
/note="24 copies 2 mer ta 75% conserved"
repeat_region 15807..16085
/note="AluSq repeat: matches 4..283 of consensus"
repeat_region 16100..16618
/note="L1ME repeat: matches 5274..5818 of consensus"
repeat_region 16649..16905
/note="AluSq repeat: matches 3..260 of consensus"
repeat_region 16918..17059
/note="L1ME repeat: matches 5100..5255 of consensus"
repeat_region 17302..17674
/note="TH1C repeat: matches 1..371 of consensus"
repeat_region 17675..19219
/note="TH1C-internal repeat: matches 1..1580 of consensus"
repeat_region 19220..19591
/note="TH1C repeat: matches 1..371 of consensus"
repeat_region 19627..19904
/note="L1M4 repeat: matches 4549..4846 of consensus"
repeat_region 19930..20313
/note="TH1C repeat: matches 1..368 of consensus"
repeat_region 21937..22041
/note="FLAM A repeat: matches 42..139 of consensus"
repeat_region 22455..22746
/note="AluX repeat: matches 1..295 of consensus"
repeat_region 23088..23137
/note="25 copies 2 mer aa 84% conserved"
repeat_region 24188..24565
/note="TH1C repeat: matches 1..371 of consensus"
repeat_region 25852..26154
/note="HAU1 repeat: matches 475..784 of consensus"
repeat_region 26159..26513
/note="L2 repeat: matches 1987..2350 of consensus"
repeat_region 27133..27321
/note="WER58A repeat: matches 27..215 of consensus"
repeat_region 27444..27537
/note="HAU1 repeat: matches 674..770 of consensus"
repeat_region 27912..28214
/note="L1M4 repeat: matches 1036..1350 of consensus"
repeat_region 29177..33723
/note="L1 repeat: matches 725..5349 of consensus"
repeat_region 37281..37600
/note="WER33 repeat: matches 1..315 of consensus"
repeat_region 38931..39055
/note="L2 repeat: matches 2580..2710 of consensus"
repeat_region 39776..40245
/note="WER1A repeat: matches 60..527 of consensus"
repeat_region 40246..41202
/note="Tigger3b repeat: matches 224..1230 of consensus"
repeat_region 42363..42506
/note="MIR repeat: matches 8..165 of consensus"
repeat_region 43820..43979
/note="MIR repeat: matches 8..157 of consensus"
repeat_region 43980..44493
/note="LTR40a repeat: matches 3..518 of consensus"
repeat_region 44494..44595
/note="MIR repeat: matches 157..261 of consensus"
misc_feature complement(45748..45750)
/note="Sequence from uni-directional primer reads only."
repeat_region 45943..46253
/note="AluX repeat: matches 1..310 of consensus"
repeat_region 46618..46876
/note="L2 repeat: matches 2433..2744 of consensus"
repeat_region 48916..49048
/note="WER86 repeat: matches 24..153 of consensus"
repeat_region 49255..49644
/note="WSTD repeat: matches 1..394 of consensus"
repeat_region 49645..51316
/note="WSTD-internal repeat: matches 1..1651 of consensus"
repeat_region 51317..51709
/note="WSTD repeat: matches 1..394 of consensus"
```

```
repeat_region 52199..52475
/note="TH1C repeat: matches 94..371 of consensus"
repeat_region 52476..52730
/note="TIGER1 repeat: matches 2078..2338 of consensus"
repeat_region 52774..53352
/note="TIGER1 repeat: matches 1702..2087 of consensus"
repeat_region 53151..53630
/note="TIGER1 repeat: matches 2..484 of consensus"
repeat_region 53953..54262
/note="AluJb repeat: matches 3..303 of consensus"
repeat_region 54608..54785
/note="89 copies 2 mer aa 57% conserved"
repeat_region 54901..55217
/note="L1M4 repeat: matches 2400..2728 of consensus"
repeat_region 55221..55414
/note="WER3 repeat: matches 3..209 of consensus"
repeat_region 55754..55952
/note="L1M1 repeat: matches 5694..5895 of consensus"
repeat_region 56125..56439
/note="AluY repeat: matches 1..311 of consensus"
repeat_region 56685..56980
/note="AluX repeat: matches 1..296 of consensus"
repeat_region 57012..57287
/note="AluJc repeat: matches 1..285 of consensus"
repeat_region 60017..60050
/note="17 copies 2 mer aa 91% conserved"
repeat_region 61032..61097
/note="33 copies 2 mer ct 71% conserved"
repeat_region 61495..61869
/note="WER47A repeat: matches 1..366 of consensus"
repeat_region 62920..63098
/note="MLTIF repeat: matches 339..540 of consensus"
repeat_region 63117..63220
/note="MLTIC repeat: matches 1..106 of consensus"
repeat_region 63289..63564
/note="MLTIF repeat: matches 29..292 of consensus"
repeat_region 64592..65042
/note="MLTIC repeat: matches 6..394 of consensus"
repeat_region 65287..65563
/note="WSTD repeat: matches 1..394 of consensus"
repeat_region 66238..66349
/note="MIR repeat: matches 28..148 of consensus"
repeat_region 66953..67403
/note="MLTIC repeat: matches 1..460 of consensus"
Query Match 18.1%; Score 43; DB 9; Length 143715;
Best Local Similarity 51.3%; Pred. No. 5.8;
Matches 100; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
```

```
QY 43 ATCAAGATAAGCGTAGATGAGTTGGGAAAAAGTGTAGAACATATCGTAAGATGAA 102
DB 29693 ATGTATTATAGCTAATGATTTTCACAGAGTCCCAAGACATTTATTCGAGAAGAA 29634
QY 103 CGTTATTGTTGAAAAATCTACTGTTACTTCAATCAGGAAACATGAAAGTAGG 162
DB 29633 CAGTCTCTTAATAATGTACCGGAAAAATTTGAATATTCATATACAGAAAGAACT 29574
QY 163 AAGTTGATTATCAGTTGATAGTTAGATTTCCTGGAACCTTTTGGAGGGATTCTATGA 222
DB 29573 AGATTACTCTCTACATATACAAAAATTAACTCAAGTGCCTTAAGACTTAATCT 29514
QY 223 AAAACTGGAAGCAAT 237
DB 29513 AAGACCTGAACTAT 29499
RESULT 14
AC021418 161828 bp DNA linear HTG 08-APR-2000
LOCUS Homo sapiens chromosome 4 clone RP11-4E4 map 4, WORKING DRAFT
DEFINITION SEQUENCE, 7 unordered pieces.
ACCESSION AC021418
VERSION AC021418.3 GI:7528069
```

HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161828)
Biren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 4, clone RP11-4E4
Unpublished
2 (bases 1 to 161828)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bede,F.,
Bohuslavsky,I., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Garayna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McSwan,P., McGurk,A., McKernan,K.,
McPheters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,P., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu.X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 8, 2000 this sequence version replaced gi:6899731.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Smit, A.F.A. & Green, P. (1996-1997).
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L4929
Center clone name: 4_E4
Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157326 bases at least Q40
Consensus quality: 159536 bases at least Q30
Consensus quality: 160510 bases at least Q20
Insert size: 194000; agarose-fp
Quality coverage: 4.5 in Q20 bases; sum-of-contigs
Quality coverage: 5.4 in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

	1	4366:	contig of 4366 bp in length
*	4367	4466:	gap of 100 bp
*	4467	16307:	contig of 11841 bp in length
*	16308	16407:	gap of 100 bp
*	16408	22925:	contig of 13518 bp in length
*	22926	30025:	gap of 100 bp
*	30026	46824:	contig of 16799 bp in length
*	46825	46924:	gap of 100 bp
*	46925	63464:	contig of 16540 bp in length
*	63465	83564:	gap of 100 bp
*	83565	83798:	contig of 20234 bp in length

FEATURES
source
1..161828
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-4E4"
/clone.lib="RPCI-11 Human Male BAC"
misc_feature
1..4366
/note="assembly_fragment"
misc_feature
4467..16307
/note="assembly_fragment"
misc_feature
16408..22925
/note="assembly_fragment"
vector_side:left
clone end:SP6
vector_side:right
30026..46824
/note="assembly_fragment"
misc_feature
46925..63464
/note="assembly_fragment"
vector_side:right
63565..83798
/note="assembly_fragment"
misc_feature
83959..161828
/note="assembly_fragment"
BASE COUNT 50687 a 27451 c 28807 g 54283 t 600 others
ORIGIN
Query Match 18.1%; Score 43; DB 2; Length 161828;
Best Local Similarity 51.3%; Pred.No. 5.7; Indels 0; Gaps 0;
Matches 100; Conservative 0; Mismatches 95;
Qy 43 ATGAAGATAAGCGTAGATCAGTTGGGAAAAAAGTGTTAGAACATATCGGTAAAGAATGAA 102
Db 20586 ATGTATTATTAGCTAACTCATTTTCACAGAGTCACCAAGAACATTTATTGGAGAAAGAA 20645
Qy 103 CCCTTATTGTTGAAAATCTACTGTTTACTTCAATCAGGGAAAAACATGAGAGTAGG 162
Db 20646 CAGTCTCTTAATAATGTCGCCGAAAAAATTAATTCATATACAGAAAGAGAAACT 20705
Qy 163 AAGTTGATTATCAGTTGATAGAGTTAGATTTTTCTGGAACITTTTGTGAGGATTTCTATGA 222
Db 20706 AGATTACTCTCTCACTATATACAAAAATTAACCAAAGTCCTTAAGACTTTAAATGT 20765
Qy 223 AAAACTGGAGCAAT 237
Db 20766 AAGACCTGAACCTAT 20780
RESULT 15
AL450123 204122 bp DNA linear HTG 13-JUN-2001
LOCUS Homo sapiens chromosome 6 clone RP11-292W23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 18 unordered pieces.
ACCESSION AL450123
VERSION AL450123.2 GI:11229253
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Burton,J.
Direct Submission
TITLE Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CS10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:11190634.
COMMENT ----- Genome Center

TITLE
JOURNAL
COMMENT

Center: Sanger Centre

Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

----- Project Information
Center project name: BA292M23

----- Summary Statistics

Assembly program: XGAP4; version 4.5
Sequencing vector: Plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 194147 bases at least Q40
Consensus quality: 198595 bases at least Q30
Consensus quality: 200843 bases at least Q20
Insert size: 202422; sum-of-contigs
Insert size: 162422; 10.6% error; agarose-fp
Quality coverage: 5.72x in Q20 bases; sum-of-contigs Quality
coverage: 8.43x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1 33126: contig of 33126 bp in length
* 33127 33226: gap of 100 bp
* 33227 114642: contig of 81416 bp in length
* 114643 114743: gap of 100 bp
* 114743 123407: contig of 8665 bp in length
* 123408 123507: gap of 100 bp
* 123508 163981: contig of 40474 bp in length
* 163982 164081: gap of 100 bp
* 164082 166908: contig of 2827 bp in length
* 166909 167008: gap of 100 bp
* 167009 169244: contig of 2236 bp in length
* 169245 172581: contig of 3237 bp in length
* 172582 172681: gap of 100 bp
* 172682 176172: contig of 3491 bp in length
* 176173 176272: gap of 100 bp
* 176273 179065: contig of 2793 bp in length
* 179066 182360: contig of 3195 bp in length
* 182361 182460: gap of 100 bp
* 182461 185537: contig of 3077 bp in length
* 185538 185637: gap of 100 bp
* 185638 188044: contig of 2407 bp in length
* 188045 188144: gap of 100 bp
* 188145 191827: contig of 3683 bp in length
* 191828 191928: gap of 100 bp
* 191929 194758: contig of 2831 bp in length
* 194759 194858: gap of 100 bp
* 194859 197211: contig of 2353 bp in length
* 197212 197311: gap of 100 bp
* 197312 199603: contig of 2292 bp in length
* 199604 199703: gap of 100 bp
* 199704 201839: contig of 2136 bp in length
* 201840 201939: gap of 100 bp
* 201940 204122: contig of 2183 bp in length.
Location/Qualifiers
1...204122
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-292M23"
/clone_lib="RPCI-11.2"
1...33126
/note="assembly fragment:01290
fragment_chain:1"
33227..114642

```

FEATURES

misc_feature
misc_feature

```

/note="assembly fragment:00700
fragment_chain:1"
114743..123407
/note="assembly fragment:02136
fragment_chain:2"
123508..163981
/note="assembly fragment:03042
fragment_chain:2"
164082..166908
/note="assembly fragment:02528
fragment_chain:3"
167009..169244
/note="assembly fragment:01791
fragment_chain:3"
169345..172581
/note="assembly fragment:00004"
172682..176172
/note="assembly fragment:00332"
176273..179065
/note="assembly fragment:00401"
179166..182360
/note="assembly fragment:00447"
182461..185537
/note="assembly fragment:00980"
185638..188044
/note="assembly fragment:01053"
188145..191827
/note="assembly fragment:02817"
191928..194758
/note="assembly fragment:02913"
194859..197211
/note="assembly fragment:03055"
197312..199603
/note="assembly fragment:03127"
199704..201839
/note="assembly fragment:03282"
201940..204122
/note="assembly fragment:03366"
1708 others
BASE COUNT 64023 a 36470 c 37269 g 64652 t
ORIGIN

```

Query Match 18.1%; Score 43; DB 2; Length 204122;
Best Local Similarity 51.3%; Pred. No. 5.5;
Matches 100; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

QY 43 ATCAAGATAGCGTAGATGAGTTGGGAAAAAAGTGTAGACATATCGTAAAGATGAA 102
Db |||||
Db 7545 ATGTTATTAAGTAACTGATTTTCACAGAGTCCAGAAACATTTATTGGAGAAAGAA 7604
QY 103 CCGTTATTTGAAAAAATCTACTGGTTAACTTCAATCAGGGAAAAACATGAAGATTAGG 162
Db |||||
Db 7605 CAGTCTTCTTAATAATGTTACCGGAAAAAATGAAATATTCATATACAGAAAGAAACT 7664
QY 163 AAGTTGATTATCAGTTGATAGATTGATTTCTGGAACTTTTGTGAGGGATTCTATGA 222
Db |||||
Db 7665 AGATTACTCTCTCTCACTATATACAAAAATTAACCTAAAGTGGCTTAAAGACTTAAATGT 7724
QY 223 AAAACTGGAAGCAAT 237
Db |||||
Db 7725 AAGACCTGAAACTAT 7739

```

Search completed: November 15, 2003, 03:34:16
Job time : 1057.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 153.606 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-96

Perfect score: 238

Sequence: 1 AATTGATACGTTGGAAA.....ATGMAAACTGGAGCAATT 238

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238	100.0	238	19 AAV03604	Neisseria meningitidis
2	233.2	98.0	16677	21 AAA1519	N. meningitidis pa
3	233.2	98.0	349980	21 AAF21612	Neisseria meningitidis
4	233.2	98.0	837096	21 AAA81489	N. meningitidis pa
5	206.8	86.9	1155	21 AAA15302	DNA encoding a pol
6	203.6	85.5	1155	21 AAA15325	DNA encoding a pol
7	200.6	84.3	1152	25 AB238717	N. gonorrhoeae nuc
8	43	18.1	11670	24 ABL54326	Chemically treated

9	40	16.8	20933	24	ABQ67123	Human angiogenesis
10	40	16.8	1664976	19	AAV21209	Methanococcus jann
11	39.8	16.7	6754	24	ABL70346	Chemically treated
12	39.8	16.7	6754	24	AA561305	Human gene regulat
13	39.6	16.6	12592	24	AA561101	Human gene regulat
14	39.2	16.5	6077	24	ABL34218	Human immune syste
15	37.4	15.7	11398	24	ABL34438	Human immune syste
16	36.8	15.5	375	22	ABL25234	Human breast cance
17	36.6	15.4	549	24	ABL78906	Human ovarian can
18	36.4	15.3	7035	22	AA546690	Tumour suppressor
19	36.2	15.2	5928	22	AA546766	Tumour suppressor
20	36.2	15.2	5928	22	ABL34239	Human immune syste
21	36.2	15.2	11260	22	AA545315	Chemically pretrea
22	36.2	15.2	11260	24	ABN80039	Human chemically m
23	36.2	15.2	11260	24	ABK28154	DNA transcription
24	36	15.1	14112	22	AA546478	Tumour suppressor
25	36	15.1	14112	24	ABQ66980	Human angiogenesis
26	36	15.1	14112	24	ABL33443	Human immune syste
27	36	15.1	14112	24	ABK31333	Signal transductio
28	36	15.1	2365589	24	ABA90521	Genomic sequence o
29	35.8	15.0	591	22	AAF82285	Staphylococcus aur
30	35.8	15.0	591	22	AA83221	GTP-binding protei
31	35.8	15.0	617	24	ABK62224	Rat sequence diffe
32	35.8	15.0	1146	24	AA14133	Plasmodium falcipa
33	35.8	15.0	2874	21	AA70144	Plasmodium falcipa
34	35.8	15.0	3107	24	ABK15041	Plasmodium cDNA en
35	35.8	15.0	6124	11	AAQ03568	Sequence encoding
36	35.8	15.0	6124	24	ABK15042	Plasmodium gene fo
37	35.8	15.0	1082138	21	AAF22305	Arabidopsis thalia
38	35.6	15.0	1440	19	AA14319	H. pylori GHPO 128
39	35.6	15.0	123219	23	AAH88703	Human DNA sequence
40	35.4	14.9	1938	20	AAZ221100	Human secreted pro
41	35.4	14.9	6029	24	ABL33992	Human immune syste
42	35.4	14.9	6954	24	ABL33391	Human immune syste
43	35.4	14.9	11398	24	ABL34439	Human immune syste
44	35	14.7	5986	24	AA561432	Human gene regulat
45	35	14.7	5986	24	ABK31498	Signal transductio

ALIGNMENTS

RESULT 1

AAV03604

ID AAV03604 standard; DNA; 238 BP.

XX AAV03604;

XX AAV03604;

XX 22-OCT-1998 (first entry)

XX Neisseria meningitidis DNA sequence B87.

XX N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;

XX region 3; pathogenicity; blood-brain barrier; diagnosis; infection;

XX meningitis; ss.

XX Neisseria meningitidis.

OS Neisseria meningitidis.

PN WO9802547-A2.

XX 22-JAN-1998.

XX 11-JUL-1997; 97WO-FR01295.

XX 12-JUL-1996; 96FR-0008768.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX (PLAC) MAX PLANCK GSS FOERDERUNG WISSENSCHAFTEN.

XX (SMIK) SMITHKLINE BEECHAM.

XX Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;

PI Vinals C;

DR WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX Example 4; Page 129; 150pp; French.

PS AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, fliA or C,

CC opc, porA, rotanase, sequence IC1106, iGA protease, pillin, pilC,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC lactamica, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX Sequence 238 BP; 85 A; 18 C; 64 G; 71 T; 0 other;

SQ

Query Match 100.0%; Score 238; DB 19; Length 238;

Best Local Similarity 100.0%; Pred. No. 1.6e-52;

Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTGGATACGTTGGAAAGGGATATTGATGGGAATGGATGAAGATAACGCTAGAT 60

DB 1 AATTGGATACGTTGGAAAGGGATATTGATGGGAATGGATGAAGATAACGCTAGAT 60

QY 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGTGAAGTAACCGTTATTGTTGAAAAAT 120

DB 61 GAGTTGGGAAAAAAGCTTTAGAACATATCGTGAAGTAACCGTTATTGTTGAAAAAT 120

QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 180

DB 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 180

QY 181 ATAGAGTTAGATTTCTGGAACCTTTGTGAGGGATCTATGAAAACTGGAAGCAATT 238

DB 181 ATAGAGTTAGATTTCTGGAACCTTTGTGAGGGATCTATGAAAACTGGAAGCAATT 238

RESULT 2

AA81519/c

ID AAA81519 standard; DNA; 16677 BP.

XX AAA81519;

AC

XX 04-DEC-2000 (first entry)

DT

XX *N. meningitidis* partial DNA sequence gnm_66 SEQ ID NO:66.

DE

XX *Neisseria meningitidis*.

OS

XX *Neisseria meningitidis*.

FN WO200022430-A2.

XX

PD 20-APR-2000.

XX

XX 08-OCT-1999; 99WO-US23573.

PF

XX 09-OCT-1998; 98US-0103794.

PR

XX 30-APR-1999; 99US-0132068.

XX

PA (CHIR) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC; Scarlato V;

PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M,

PI Rappuoli R, Pizza M;

XX WPI; 2000-318079/27.

XX

PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be

PT used in the diagnosis and treatment of *N. meningitidis* infection and

PT other *Neisseria* infections, for example, *N. gonorrhoeae* -

XX Claim 7; Page 1406-1411; 1760pp; English.

XX

CC The present invention describes methods of obtaining immunogenic

CC proteins from *Neisseria* genomic sequences. AAA81433 to AAA82414

CC represent specifically claimed *Neisseria meningitidis* genomic DNA

CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to

CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81322 to

CC AAA81432 represent *Neisseria meningitidis* MenB polynucleotide ORF

CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The

CC composition can be used as a medicament (or in the manufacture of a

CC medicament) for treating, preventing or diagnosing infection due to

CC *Neisseria* bacteria. For example, some of the identified proteins could

CC be components of vaccines against *Neisseria meningitidis* B; against all serotypes;

CC and/or against all pathogenic *Neisseria*. Identification of sequences

CC from the bacterium will also facilitate production of biological probes,

CC particularly organism-specific probes. Attempts to make efficacious

CC *Meningococcus B* vaccines have failed mainly due to antigen tolerance.

CC Multivalent vaccines have also been tried but none have successfully

CC overcome antigenic variability. The provision of further, complete

CC sequences may provide an opportunity to identify secreted or surface

CC exposed proteins that may be presumed targets for the immune system and

CC which are not antigenically variable or at least more conserved than

CC other more variable regions.

XX Sequence 16677 BP; 4119 A; 4509 C; 3979 G; 4070 T; 0 other;

SQ

Query Match 98.0%; Score 233.2; DB 21; Length 16677;

Best Local Similarity 98.7%; Pred. No. 7.2e-51;

Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATTGGATACGTTGGAAAGGGATATTGATGGGAATGGATGAAGATAACGCTAGAT 60

DB 5014 AATTGGATACGTTGGAAAGGGATATTGATGGGAATGGATGAAGATAACGCTAGAT 4955

QY 61 GAGTTGGGAAAAAAGTGTGAAACATATCGTGAAGTAACCGTTATTGTTGAAAAAT 120

DB 4954 GAGTTGGGAAAAAAGTGTGAAACATATCGTGAAGTAACCGTTATTGTTGAAAAAT 4895

QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 180

DB 4894 CTACTGGTTAACTTCAATCAGGAAAAATATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 4835

QY 181 ATAGAGTTAGATTTCTGGAACCTTTGTGAGGGATCTATGAAAACTGGAAGCAATT 238

DB 4834 ATAGAGTTAGATTTCTGGAACCTTTGTGAGGGATCTATGAAAACTGGAAGCAATT 4777

RESULT 3

AAF21612/c

ID AAF21612 standard; DNA; 349980 BP.

XX

XX AAF21612;

XX

DT 13-MAR-2001 (first entry)

DE

XX *Neisseria meningitidis* B nucleotide sequence SEQ ID NO:113.

XX

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;

diagnosis; antigen; detection; infection; gene therapy; antibacterial; ds.

Neisseria meningitidis.

WO200066791-A1.

09-NOV-2000.

08-MAR-2000; 2000WO-US05928.

30-APR-1999; 99US-0132068.

08-OCT-1999; 99WO-US23573.

28-FEB-2000; 2000GB-0004695.

(CHIR) CHIRON CORP.

(GENO-) INST GENOMIC RES.

Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Massignani V; Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R; Frazer CM, Grandi G; WPI; 2000-647603/62.

Neisseria meningitidis B full length genome sequence and open reading frames are used to detect, treat and prevent Neisserial infections - Claim 7; Appendix A; 692pp; English.

The present invention describes the full length genome of Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613 represent fragments of the NMB genomic sequence, as the sequence was too long to go in a record on its own it was split into 8 sequences which overlap each other at the beginning and end of each sequence by 4980 bp (i.e. the last 4980 bp of AAF21544 is repeated at the beginning of AAF21607, the last 4980 bp of AAF21607 are repeated at the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins given in AAF58550 to AAF58593, and AAF21589 to AAF21606 represent PCR primers which are used in the exemplification of the present invention. The NMB genome and fragments from it have antibacterial activity, and can be used in vaccines and gene therapy. Neisseria nucleic acids, proteins and/or antibodies which binds to the proteins can be used in compositions for treating or preventing infection due to Neisserial bacteria or as a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised to Neisserial bacteria. Computers, computer memory, computer storage medium or computer databases can be used in a search to identify open reading frames (ORFs) or coding sequences within the NMB genome. The DNA sequences provide further opportunities to find antigenic or immunogenic proteins which are more effective in vaccines than the outer membrane proteins currently used.

Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

Query Match 98.0%; Score 233.2; DB 21; Length 349980;

Best Local Similarity 98.7%; Pred. No. 1.4e-50;

Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATTGGATACGTTGGAAAGCGATATTGATTCGGATGGATGAAGATGAAGCGTAGAT 60

72678 AATTGGATACGTTGGAAAGCGATATTGATTCGGATGGATGAAGATGAAGCGTAGAT 72619

61 GAGTTGGGAAAAAGTGTAGACATATCGTGAAGATGAACCGTTATTGTTGAAAAAT 120

72618 GAGTTGGGAAAAAGTGTAGACATATCGTGAAGATGAACCGTTATTGTTGAAAAAT 72559

121 CTACTGGTTAACTTCAATCAGGGAACATGACGAAGTTAGGATGATTATTCAGTTG 180

72558 CTACTGGTTAACTTCAATCAGGGAACATGACGAAGTTAGGATGATTATTCAGTTG 72499

181 ATAGAGTTAGATTTTCTGGAACTTTGTGAGGGATTCATGAAAACTGGAAGCAATT 238

72498 ATAGAGTTAGATTTTCTGGAACTTTGTGAGGGATTCATGAAAACTGGAAGCAATT 72441

RESULT 4

AAA81489/C

ID AAA81489 standard; DNA; 837096 BP.

XX AAA81489;

AC AAA81489;

XX 04-DEC-2000 (first entry)

DT 04-DEC-2000 (first entry)

XX N. meningitidis partial DNA sequence grm_37 SEQ ID NO:37.

DE Neisseria meningitidis; Neisseria gonorrhoeae; Genome; immunogenic;

XX Neisseria meningitidis; Neisseria gonorrhoeae; Genome; immunogenic;

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

KW Meningococcus B; MenB; ds.

XX Neisseria meningitidis.

OS Neisseria meningitidis.

XX WO200022430-A2.

PN WO200022430-A2.

XX 20-APR-2000.

PD 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

XX 09-OCT-1998; 98US-0103794.

XX 30-APR-1999; 99US-0132068.

PR 30-APR-1999; 99US-0132068.

XX (CHIR) CHIRON CORP.

XX (CHIR) CHIRON CORP.

XX (CHIR) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

XX Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

XX Rappuoli R, Pizza M;

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

XX WPI; 2000-318079/27.

Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Query Match 98.0%; Score 233.2; DB 21; Length 837096;

Best Local Similarity 98.7%; Pred. No. 1.7e-50;

Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 60
 DB 437449 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 437390
 QY 61 GAGTTGGGAAAAAAGTGTAGAACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 120
 DB 437389 GAGTTGGGAAAAAAGTGTAGAACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 437330
 QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTATCAGTTG 180
 DB 437329 CTACTGGTTAACTTCAATCAGGAAAAATATGAAGAAGTTAGGAAGTTGATTATCAGTTG 437270
 QY 181 ATAGAGTTAGATTCTCGAACTTTCTGAGGATTTCTATCAAAACTGGAAGCAATT 238
 DB 437269 ATAGAGTTAGATTCTCGAAATTTTGTGAGGATTTCTATGAAAAAACTGGAAGCAATT 437212

RESULT 5
 AAA15302
 ID AAA15302 standard; DNA; 1155 BP.
 AC AAA15302;
 DT 04-SEP-2000 (first entry)
 DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.
 XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 OS *Neisseria meningitidis*.
 FH Key Location/Qualifiers
 FT CDS 1..1155
 FT /*tag= a
 XX WO200026375-A2.
 PD 11-MAY-2000.
 PF 28-OCT-1999; 99WO-FR02643.
 PR 30-OCT-1998; 98FR-0013693.
 XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 DR WPI; 2000-365622/31.
 DR P-PSDB; AAY93271.
 PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic or preventative vaccines and for diagnosis
 PS Claim 2; Page 73-75; 187pp; French.
 XX The present sequence encodes a protein that is specific for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides, or their antigenic fragments, are used in vaccines to treat or protect against *Neisseria* infections, particularly by *N. meningitidis*. The polynucleotide sequence is also used for recombinant production of *Neisseria* and to produce attenuated *Neisseria* strains that overexpress it, or express it in a non-toxic mutant form.
 XX Sequence 1155 BP; 335 A; 269 C; 248 G; 303 T; 0 other;

Query Match 86.9%; Score 206.8; DB 21; Length 1155;
 Best Local Similarity 99.0%; Pred. NO. 2.9e-44;
 Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 60
 DB 437449 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 437390

DB 946 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 1005
 QY 61 GAGTTGGGAAAAAAGTGTAGAACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 120
 DB 1006 GAAATTGGGAAAAAAGTGTAGAACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 1065
 QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTATCAGTTG 180
 DB 1066 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTATCAGTTG 1125
 QY 181 ATAGAGTTAGATTCTCGAACTTTTGTGA 210
 DB 1126 ATAGAGTTAGATTCTCGAACTTTTGTGA 1155

RESULT 6
 AAA15325
 ID AAA15325 standard; DNA; 1155 BP.
 AC AAA15325;
 DT 04-SEP-2000 (first entry)
 DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.
 XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.
 OS *Neisseria gonorrhoea*.
 FH Key Location/Qualifiers
 FT CDS 1..1155
 FT /*tag= a
 XX WO200026375-A2.
 PD 11-MAY-2000.
 PF 28-OCT-1999; 99WO-FR02643.
 PR 30-OCT-1998; 98FR-0013693.
 XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
 PI Tinsley C, Perrin A;
 DR WPI; 2000-365622/31.
 DR P-PSDB; AAY93295.
 PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic or preventative vaccines and for diagnosis
 PS Claim 4; Page 144-146; 187pp; French.
 XX The present sequence encodes a protein that is specific for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides, or their antigenic fragments, are used in vaccines to treat or protect against *Neisseria* infections, particularly by *N. meningitidis*. The polynucleotide sequence is also used for recombinant production of *Neisseria* and to produce attenuated *Neisseria* strains that overexpress it, or express it in a non-toxic mutant form.
 XX Sequence 1155 BP; 337 A; 265 C; 246 G; 307 T; 0 other;

Query Match 85.5%; Score 203.6; DB 21; Length 1155;
 Best Local Similarity 98.1%; Pred. NO. 1.9e-43;
 Matches 206; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 60
 DB 946 AATTGGATACGTTGGAAAAGGGATATTGATTGGGAATGGGATGAAGATAAGCGTAGAT 1005

QY 61 GAGTTGGGAAAAAGTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 1006 GAGTTGGGAAAAAGTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 1065
QY 121 CTACTGGTTAACTTCAATCAGGAAACATGAAGAGTTAGAAAGTTGATTTATCAGTTG 180
Db 1066 CTACTGGTTAACTTCAATCAGGAAACATGAAGAGTTAGAAAGTTGATTTATCAGTTG 1125
QY 181 ATAGAGTTAGATTTCTGGAACCTTTTGTGA 210
Db 1126 ATAGAGTTAGATTTCTGGAATTTTGTGA 1155

RESULT 7
ABZ38717
ID ABZ38717 standard; DNA; 1152 BP.
XX AC ABZ38717;
XX DT 07-MAR-2003 (first entry)
XX N. gonorrhoeae nucleotide sequence SEQ ID 2023.
DE Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX KW Neisseria gonorrhoeae.
XX OS
XX PN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB02069.
XX PR 12-FEB-2001; 2001GB-0003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
DR P-PSDB; ABP77747.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection
XX Disclosure; Page 334-335; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention.
XX
SQ Sequence 1152 BP; 336 A; 265 C; 245 G; 306 T; 0 other;

Query Match 84.3%; Score 200.6; DB 25; Length 1152;
Best Local Similarity 98.1%; Pred. No. 1.2e-42;
Matches 203; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AATTGGGATACGTTGGAAAGGGATATTGATTGGGAATGGATGAAGATGAAGCGTAGAT 60
Db 946 AATTGGGATACGTTGGAAAGGGATATTGATTGGGAATGGATGAAGATGAAGCGTAGAT 1005
QY 61 GAGTTGGGAAAAAGTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 1006 GAGTTGGGAAAAAGTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAAT 1065
QY 121 CTACTGGTTAACTTCAATCAGGAAACATGAAGAGTTAGAAAGTTGATTTATCAGTTG 180
Db 1066 CTACTGGTTAACTTCAATCAGGAAACATGAAGAGTTAGAAAGTTGATTTATCAGTTG 1125

QY 181 ATAGAGTTAGATTTCTGGAACCTTTG 207
Db 1126 ATAGAGTTAGATTTCTGGAATTTG 1152
RESULT 8
ABL54326
ID ABL54326 standard; DNA; 11670 BP.
XX AC ABL54326;
XX DT 29-JUL-2002 (first entry)
XX Chemically treated apoptosis gene complementary to gene #13.
XX Apoptosis; HIV; Bloom syndrome; cardiopathy;
KW neurodegenerative disorder; Herpes simplex virus; renal ischaemia;
KW amyotrophic lateral sclerosis; cancer; ds.
XX OS Unidentified.
XX PN WO200177164-A2.
XX PD 18-OCT-2001.
XX PF 06-APR-2001; 2001WO-EP03969.
XX PR 06-APR-2000; 2000DE-1019058.
XX PR 07-APR-2000; 2000DE-1019173.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-017444/02.
XX Chemically modified sequences of genes associated with apoptosis are
PT useful to determine methylation patterns of genomic DNA samples for
PT diagnosis of associated diseases such as cancer
XX Claim 1; Seq ID #26; 24pp; English.
XX This invention relates to chemically pre-treated DNA of genes
CC associated with apoptosis. The nucleic acids are used to allocate
CC patients for specific therapy for HIV infection, Bloom syndrome,
CC cardiopathy, aging, neurodegenerative disorders, Herpes simplex, virus
CC infection, renal ischaemia, amyotrophic lateral sclerosis, solid tumours
CC and cancers. This nucleotide sequence represents a chemically
CC treated apoptosis gene. Even SEQ ID numbers are the complementary
CC DNA strands to the odd SEQ ID numbers. The sequence data for this
CC patent is not represented in the printed specification but is based on
CC information supplied by the European patent office.
XX
SQ Sequence 11670 BP; 3563 A; 238 C; 2443 G; 5426 T; 0 other;

Query Match 18.1%; Score 43; DB 24; Length 11670;
Best Local Similarity 49.3%; Pred. No. 0.15;
Matches 112; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 3 TTGGGATACGTTGGAAAGGGATATTGATTGGGAATGGATGAAGATGAAGCGTAGATGA 62
Db 265 TTAGGTTAGTTGAAATATATTTTATTTATGATTATTTATTTTGAAGTGGG 324
QY 63 GTTGGGAAAAAGTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAATCT 122
Db 325 GTAAGAAATAAAGTGTGTTTATTATAGATTATTAATTAATTTTCTTTTGGATA 384
QY 123 ACTGGTTAACTTCAATCAGGAAACATGAAGAGTTAGAAAGTTGATTTATCAGTTGAT 182
Db 385 GAGTTGATTTTAGATCGTAAATATAGTAAATATTAATATTTTAAATTTATGATTAGT 444

QY	183	AGAGTTAGATTTCTGGAACCTTTTGAGGAGTCTATGAAAACTG	229
DB	445	AGAGTTAGGTCGTGTTGTTTATATATTTTATTTTATAGTTG	491
RESULT 9			
ID	ABQ67123	standard; DNA; 20933 BP.	
XX	AC	ABQ67123;	
XX	DT	28-AUG-2002 (first entry)	
XX	DE	Human angiogenesis associated polynucleotide SEQ ID NO 153.	
XX	XX	Human; angiogenesis; methylation; eye disease; glaucoma; tumour;	
KW	inflammation; rheumatoid arthritis; diabetic retinopathy; antitubercu-		
KW	macular degeneration; inflammatory bowel disease; Crohn's disease;		
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;		
KW	antiarteriosclerotic; ds.		
XX	OS	Homo sapiens.	
XX	XX	WO200246454-A2.	
XX	PN	13-JUN-2002.	
XX	PD	06-DEC-2001; 2001WO-EP14320.	
XX	PF	06-DEC-2000; 2000DE-1061335.	
XX	PR	(EPIG-) EPIGENOMICS AG.	
XX	PA	Schacht O;	
XX	PI	WPI; 2002-500450/53.	
XX	DR	New nucleic acid fragments from chemically treated	
XX	PT	angiogenesis-associated genes, useful for determining methylation	
XX	PT	status, e.g. in diagnosis or treatment of cancer	
XX	PS	Claim 1; SEQ ID NO 153; 41pp + Sequence Listing; German.	
XX	XX	The invention relates to a nucleic acid (I) comprising a segment of 18	
CC	bases of chemically pretreated DNA of angiogenesis-associated genes (II)		
CC	having sequences (ABQ671-ABQ67178) or their complements. (I), also		
CC	related oligomers, are used to evaluate the methylation status and/or		
CC	single-nucleotide polymorphisms, in angiogenesis-related genes, for		
CC	diagnosis and treatment of eye diseases, proliferative retinopathy,		
CC	neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,		
CC	diabetic retinopathy, macular degeneration caused by neovascularisation,		
CC	psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and		
CC	Crohn's disease.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 20933 BP; 4999 A; 396 C; 5359 G; 10179 T; 0 other;		
SQ	Query Match	16.8%; Score 40; DB 24; Length 20933;	
	Best Local Similarity	50.5%; Pred. No. 1;	
	Matches	97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;	
QY	25	TATTTTCATCGGAATCGGATAGATAGCGTATAGTGGGAAAAAGTCTAGAA	84
DB	17421	TATTTGGGAGGTAGGTAGGAGGATTTTTCAGTTTAGGATTTAGGTTTATGAGT	17480
QY	85	CATATCGGTAGAAATGAACCGTTATGTTGAAAAATCTACTGGTTAACTCAATCAGGA	144
DB	17481	CGTGACGGTCCGATGTTATTTTAGTTATAGATGAGATTTGTTTAAAGTAAAAAGA	17540
QY	145	AAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTGATAGTTATTTCTCGA	204
DB	17541	GAAGGAGATAAGTGAGGTATTTGAGTAAGTGGGTAGGATGGTTAGTATAGT	17600
QY	205	TTGTGAGGGATT	216
DB	17601	TATAGGAGGTT	17612
RESULT 10			
ID	AAV21209	standard; DNA; 1664976 BP.	
XX	AC	AAV21209;	
XX	DT	10-NOV-1998 (first entry)	
XX	DE	Methanococcus jannaschii circular chromosome.	
XX	XX	Methanococcus jannaschii; methanogenic archaeon; circular chromosome;	
KW	genome; autotrophic; extrachromosomal element; identification; ds.		
XX	OS	Methanococcus jannaschii.	
XX	XX	WO9807830-A2.	
XX	PN	26-FEB-1998.	
XX	PD	22-AUG-1997; 97WO-US14900.	
XX	PF	22-AUG-1996; 96US-0024428.	
XX	PR	(GENO-) INST GENOMIC RES.	
XX	PA	(UNII) UNIV ILLINOIS FOUND	
XX	PA	(UIJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.	
XX	PI	Bult CJ, Smith HO, Venter JC, White OR, Woese CR;	
XX	XX	WPI; 1998-169145/15.	
XX	XX	Complete genome sequence of methano-genic archaeon, Methanococcus	
XX	PT	jannaschii - useful in identification of M. jannaschii genome	
XX	PT	fragment	
XX	PS	Claim 13; Page 152-585; 614pp; English.	
XX	XX	The present sequence represents the complete 1.66-megabase pair genome	
CC	sequence of the Methanococcus jannaschii circular chromosome. The		
CC	present invention describes M. jannaschii open reading frames from the		
CC	genome sequence. The invention also describes a computer based system		
CC	for identifying fragments of the M. jannaschii genome that are		
CC	homologous to target nucleotide sequences, comprising: (a) data storage		
CC	means comprising the nucleotide sequence of the 1664976, 58407 or 16550		
CC	bp sequence (see AAV21209, AAV21210 and AAV21211), or a nucleotide		
CC	sequence at least 99.9% identical to it; (b) search means for comparing a		
CC	target sequence to the nucleotide sequence of the data storage means to		
CC	identify a homologous sequence, and (c) retrieval means for obtaining		
CC	the homologous sequence. The method, which is based on whole genome		
CC	random sequencing of an autotrophic archaeon M. jannaschii, the genome		
CC	of which consists of 3 physically distinct elements, a large circular		
CC	chromosome (the 1664976 bp sequence given in AAV21209), a large circular		
CC	extra-chromosomal element (the 58407 bp sequence given in AAV21210), and		
CC	a small circular extra-chromosomal element (the 16550 bp sequence given		
CC	in AAV21211), can be used in the identification of M. jannaschii genome		
CC	fragment.		
XX	Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other;		
SQ	Query Match	16.8%; Score 40; DB 19; Length 1664976;	
	Best Local Similarity	55.9%; Pred. No. 2.6; Mismatches 60; Indels 0; Gaps 0;	
	Matches	76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;	
QY	70	AAAAAGTCTAGACATATCGTATAGATGACCGTTATTTGAAAAATCTATGTT	129

Db 1095978 AACTAAATTTGCAATACCTGGGGAGGATAGACAAATAGTGGAGATTTATTGGTC 1096037

Qy 130 AACITCATCAGGGAACATGACAGACTTAGGAGTTGATTATCATCTTATAGAGTTA 189

Db 1096038 TATTCCTAACCTGAGAAATACAGCAAAATTTGGATTATATGAAGCATTAGATAGAGTTA 1096097

Qy 190 GATTTCTGGAACTTT 205

Db 1096098 GAGACGAGGAGCTTT 1096113

RESULT 11

ID ABL70346

XX ABL70346 standard; DNA; 6754 BP.

AC ABL70346;

XX 01-JUL-2002 (first entry)

DE Chemically treated cell signalling DNA sequence complementary to#118.

XX Cell signalling; cytosine methylation; cell signalling disease;

KW cancer; tumour; cytostatic; ds.

XX Unidentified.

OS WO200202807-A2.

PN 10-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07471.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

PA Olek A, Piepenbrock C, Berlin K;

PI WPI; 2002-154758/20.

DR Nucleic acid, useful for diagnosis and therapy of diseases associated

XX with cell signalling e.g. cancer comprises chemically modified genomic

PT sequences of genes associated with cell signalling -

XX Claim 1; SEQ ID NO 236; 24pp+sequence listing; English.

PS The invention relates to a nucleic acid comprising a sequence of at least

CC 18 bases of a segment of chemically pretreated DNA of genes associated

CC with cell signalling. The activity of the modified sequences of the

CC invention may be described as cytostatic. The object of the invention is

CC to provide the chemically modified DNA of genes associated with cell

CC signalling, as well as oligonucleotides and/or PNA-oligomers for

CC detecting cytosine methylations, as well as a method which is

CC particularly suitable for the diagnosis and/or therapy of genetic and

CC epigenetic parameters of genes associated with cell signalling. The

CC chemically modified DNA provided by the invention is useful for diagnosis

CC and therapy of diseases such as solid tumours and cancer. The sequences

CC given in records ABL70111-ABL70626 represent chemically pre-treated

CC genomic DNA's of genes associated with cell signalling.

CC Note: The sequence data for this patent is not represented in the printed

CC specification, but is based on sequence information supplied by the

CC European Patent Office.

XX Sequence 6754 BP; 1864 A; 136 C; 1369 G; 3385 T; 0 other;

SQ Query Match 16.7%; Score 39.8; DB 24; Length 6754;

Best Local Similarity 50.3%; Pred. No. 0.89;

Matches 98; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 39 TCGGATGAGATACCGTAGTACGTTGGGAAACATGTAGAACATATCGGTACAA 98

Db 6028 TAGGAGTCGAGCGGGTGGATTATTGAGTTAGGATTCGAGATTGTTGTTAATA 6087

Qy 99 TGAACCGTTATTGTTCAAAATCTACTGTTTAACTTCAATCAGGGAAACATGAAGAAGT 158

Db 6088 TGGTGGTTAGTTATAGTATAGTGGGGTTTTATTATATATAAAATATATAAAATTAG 6147

Qy 159 TAGGAAGTTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGAACTTTTCTGAGGGATTT 218

Db 6148 TTAGACGTGGTGATATATATTGTTAGTTTATTCGGGATGTTGAGATAGAGATTT 6207

Qy 219 ATGAAACAACTGGAG 233

Db 6208 ATTTAAATTTGGGAG 6222

RESULT 12

AAS61305

ID AAS61305 standard; DNA; 6754 BP.

XX AAS61305;

AC AAS61305;

XX 29-JAN-2002 (first entry)

DT Human gene regulation-associated gene oligonucleotide #260.

DE Human; Gene regulation-associated gene; severe combined immunodeficiency;

XX cardiac damage; inflammatory response; Haemophilia; Werner syndrome;

KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;

XX renal disease; Preeclampsia; cardiac allograft vascular disease;

KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;

XX immunostimulant; cardiac; antiinflammatory; coagulant; antiasthmatic;

KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX Homo sapiens.

OS WO200177375-A2.

PN 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.

XX 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

PA Olek A, Piepenbrock C, Berlin K;

PI WPI; 2002-017470/02.

DR New nucleic acid sequences from chemically modified genes associated

XX with gene regulation, useful for analysing cytosine methylations for

PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency

PT disease -

XX Disclosure; SEQ ID No 266; 26pp; English.

PS The invention relates to 224 nucleic acid sequences comprising at least

CC 18 bases of a chemically pretreated gene associated with gene regulation

CC selected from 43 known genes (or complementary sequences). The

CC chemical pretreatment converts cytosine bases unmethylated at the

CC 5-position to uracil or another base with hybridisation behaviour

CC dissimilar to cytosine, to enable analysis of cytosine methylations.

CC The DNA sequences, oligomers (or sets/arrays) and method are

CC useful in the diagnosis of diseases (or predisposition to diseases)

CC associated with gene regulation and in therapy of such diseases, by

CC enabling analysis of the cytosine methylation patterns of such genes,

CC kits are provided. They are especially useful in diagnosis

CC and therapy of e.g. severe combined immunodeficiency disease, cardiac

CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,

CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,

CC preeclampsia, graft versus-host disease. The present sequence is a

CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 6754 BP; 1864 A; 136 C; 1369 G; 3385 T; 0 other;
 Query Match 16.7%; Score 39.8; DB 24; Length 6754;
 Best Local Similarity 50.3%; Pred. No. 0.89;
 Matches 98; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
 QY 39 TGGGATGAAGATAAGCGTAGATGAGTGGGGAAGAAAGTGTAGAACATATCGGTAGAA 98
 Db 6028 TAGGAGGTCGAGCGGGTGGATTATTTGAGGTAGGAGTTCGAGATTAGTTGGTTAATA 6087
 QY 99 TGAACGGTTATGTGAAATCTACTGTTAACTTCAATCAGGGAAGAAATCAAGAGAGT 158
 Db 6088 TGGTGGTTAGTTATAGTATAGTCGGGTTTATTTATATATAAAAAATATAAAAAATTAG 6147
 QY 159 TAGGAGTGTGATTTATCAGTTGATAGATTAGATTTCTGGAATTTTGTGAGGATTC 218
 Db 6148 TTACGCTGGTGCATATATTTTGTAGTTTATTCGGGATGTTGAGATAGGAGATTT 6207
 QY 219 ATGAAAGAACTGGAAG 233
 Db 6208 ATTTAAATTTGGAG 6222
 RESULT 13
 ID AAS61101 standard; DNA; 12592 BP.
 XX AAS61101;
 AC AAS61101;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human gene regulation-associated gene oligonucleotide #56.
 KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chorzen syndrome;
 KW renal disease; freckleplasia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200177375-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03968.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-017470/02.
 XX
 XX New nucleic acid sequences from chemically modified genes associated
 XX with gene regulation, useful for analysing cytosine methylations for
 XX diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 XX disease -
 XX
 PS Claim 1; SEQ ID No 57; 26pp; English.

XX
 CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chorzen syndrome, renal disease,
 CC presenclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 12592 BP; 3783 A; 195 C; 2736 G; 5878 T; 0 other;
 Query Match 16.6%; Score 39.6; DB 24; Length 12592;
 Best Local Similarity 48.6%; Pred. No. 1.2;
 Matches 108; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
 QY 16 GAAAGGGATATTTGATTCGGGAATCGGATGAAGATAGCGTAGATGTCGGGAAAAAA 75
 Db 7892 GAAGGAGTATTTTGGTTGGGGATAAATGTTTAGGAGATGATTTATATATGGAATTTAA 7951
 QY 76 GTGTTAGAACATATCGGTAGAACATGACCGTTATTTGTGAAAAATCTACTGGTTAACTTC 135
 Db 7952 GAGTAATATATTTGGTTATTTAATGAGAAATTAGATTATGCTGTTATGTAATGAAT 8011
 QY 136 ATCAGGGAACATGAAGAAAGTTAGGAAGTTGATTATTCAGTTGATAGATTAGATTT 195
 Db 8012 AATTATTATATGAGTAGAAGAGTTTTTGGTAGAGGAATTATGTAATAAAGATTTATT 8071
 QY 136 CTGGAACCTTTTGTGAGGATTCATGAAAACTGGAAGCAAT 237
 Db 8072 TCGGAATATTTATTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8113
 RESULT 14
 ID ABL34218
 XX ABL34218;
 AC ABL34218;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 2191.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; neutropic;
 KW neuroprotective; anti-Hiv; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX Claim 1; SEQ ID NO 2191; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX Sequence: 6077 BP; 1715 A; 44 C; 1330 G; 2988 T; 0 other;
SQ
Query Match 16.5%; Score 39.2; DB 24; Length 6077;
Best Local Similarity 50.5%; Pred. No. 1.2;
Matches 95; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 24 ATATTGTTGGGATGGATGAAGATAGATGAGTGGGGAAGAGTGTAGA 83
Db 824 ATAGTGGTTTGGGTGGAAGTATTTTAATTTGGAATGTTGTGTTGGTGGATGGA 883
QY 84 ACATATCGGTGAAGATGACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGG 143
Db 884 GTATTAAATTGAGTTTAGAGTCTTTTTCAGAGTTTTTTTGTGTTAGATGATTTTGAG 943
QY 144 AARACATCAAGAGTTAGGAGTTGATTTATCATGATGATAGATTTTCTGGAACT 203
Db 944 GAATGATGATTTATTGTTGTTGTTTAAATGATGATATAGTTTATATAAT 1003
QY 204 TTTGTGAG 211
Db 1004 TATATGAG 1011
RESULT 15
ABL34438
ID ABL34438 standard; DNA; 11398 BP.
XX ABL34438;
XX AC
XX XX
XX 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 2411.
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anti-anaemic; cytostatic; neutropenic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX XX
XX Homo sapiens.
XX OS
XX WO200200928-A2.
XX PN
XX XX
XX PD
XX 03-JAN-2002.
XX XX
XX 02-JUL-2001; 2001WO-EP07537.
XX PF

XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX Claim 1; SEQ ID NO 2411; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX Sequence: 11398 BP; 2670 A; 275 C; 2835 G; 5618 T; 0 other;
SQ
Query Match 15.7%; Score 37.4; DB 24; Length 11398;
Best Local Similarity 50.9%; Pred. No. 4.2;
Matches 89; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 62 AGTTGGGAAAAAGTGTAGAACATATCGGTAAGATGAACCGTTATTGTTGAAAAATC 121
Db 2565 AGTGAGGTAGTATTTTTTTTAAATGTTATATTTTAAATTTTATATATATAT 2624
QY 122 TACTGTTTAACTTCAATCAGGGAACATGAAGAGTTAGGAAGTTGATTATCAGTTCA 181
Db 2625 TTTTGTAGTTTATTAATATATATATTAATTAATTAATTTTGTAGTTTGTAGTTT 2684
QY 182 TAGAGTTAGATTTCTCGAACTTTTGTGAGGATTTCTATGAAAAAACTGGAAGCAA 236
Db 2685 TAAATTTAGTTTATTGGATGTTTTTTTAAAGAGTAGAGATAGAGAGGAGTAA 2739
Search completed: November 15, 2003, 00:35:50
Job time : 159.606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 ; Search time 1587.72 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-96
Perfect score: 238
Sequence: 1 AATTGGATACGTTGAAAA.....ATGAAACTGGAAGCAATT 238

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hci.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47.6	20.0	1201	13	BX422549
2	45.8	19.2	1112	9	AL552404
3	44.2	18.6	989	13	BX361014
4	43.2	18.2	905	13	BX369737

C	5	43	18.1	353	28	AQ222887
	6	43	18.1	575	28	AQ506302
	7	43	18.1	1101	29	CNS00067
C	8	42.4	17.8	483	28	B53839
	9	42	17.6	721	14	CB285716
	10	41.8	17.6	1175	13	BX423299
C	11	41.6	17.5	862	28	AZ549439
C	12	41.6	17.5	1101	29	CNS00039G
C	13	41.2	17.3	641	9	AV885336
C	14	41.2	17.3	679	9	AV849915
C	15	41.2	17.3	737	13	BW092427
C	16	41.2	17.3	1083	28	BH180017
	17	41.2	17.3	1083	29	CNS07LYX
C	18	41	17.2	833	13	BQ136570
	19	40.8	17.1	1201	13	BX376686
	20	40.6	17.1	842	29	CNS000KPF
	21	40.4	17.0	1201	13	BX360626
C	22	40.4	17.0	1322	29	CC258640
C	23	40.2	16.9	1191	9	AL558073
	24	40	16.8	519	10	BE432539
C	25	40	16.8	938	29	CNS00067J
	26	40	16.8	1101	29	CNS0182P
C	27	39.4	16.6	693	28	BH517222
C	28	39.2	16.5	983	29	CNS014HQ
C	29	39.2	16.5	1201	13	BX356174
	30	39	16.4	434	28	AQ493305
	31	39	16.4	442	13	BQ104067
C	32	39	16.4	859	29	CNS004YY
	33	38.8	16.3	520	13	BQ633195
C	34	38.8	16.3	615	12	BJ405869
C	35	38.8	16.3	855	13	BX355304
C	36	38.8	16.3	1200	13	BX437740
	37	38.8	16.3	1201	13	BX461310
	38	38.6	16.2	855	28	AZ691600
	39	38.6	16.2	861	28	AZ678433
	40	38.6	16.2	971	28	BH165867
	41	38.6	16.2	1101	29	CNS000D1
C	42	38.6	16.2	1101	29	CNS017KX
C	43	38.4	16.1	1201	29	AQ440391
C	44	38.4	16.1	1201	29	CNS01651
C	45	38.2	16.1	737	13	BQ515493

ALIGNMENTS

RESULT 1
BX422549
LOCUS
DEFINITION BX422549 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CSODM007YJ05.5-PRIME, mRNA sequence.
ACCESSION BX422549
VERSION BX422549.1 GI:30769217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8205.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODM007YJ05.5-PRIME-8205.f.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODM007CB03QP1.

FEATURES

source
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM007YU05"
/tissue type="FETAL LIVER"
/dev stages="fetal"
/clone lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT 311 a 219 c 275 g 213 t 183 others
ORIGIN
Query Match 20.0%; Score 47.6; DB 13; Length 1201;
Best Local Similarity 45.0%; Pred. No. 0.29;
Matches 86; Conservative 20; Mismatches 85; Indels 0; Gaps 0;
QY 30 GATTGGGAATGGGATGAAGTAAAGCGTAGATGAGTGGGAAAAAAGTGTAGAACATAT 89
DB 648 GMTTAAAGAAAGATAGAGGAGCGGTGSAAGAGCGGGTGGAGNATAGATTWAA 707
QY 90 CGTGAAGTAATGAACGGTATTGTTGAAAAATCTACTCGTTAACTTCAATCAGGGAAAAACA 149
DB 708 KGTGAAAAAGGAGAAKWTATTAGTTATTTCGTTTASRAGTTTBGAGAGAGGAGGAASGAR 767
QY 150 TGAAGAAGTTAGGAGGTTGATTATCATGTTGATGATGATGATGATGATGATGATGATG 209
DB 768 AGAARAAGAGAGGAGCGGTTWAAATGWAAGAGGTTWAAATWAAAGATTTWTTTS 827
QY 210 AGGATTTCTAT 220
DB 828 GGGASSSCTAT 838
RESULT 2
AL552404 1112 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL552404 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI070YP06 5-PRIME, mRNA sequence.
ACCESSION
AL552404.2 GI:31274219
VERSION
EUT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1112)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
COMMENT
On Feb 15, 2001 this sequence version replaced gi:12891270.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5872.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI070H03QP1&cluster=5872.r. Contact :
Feng Liang Email : fliang@lifetech.com URL : Corporation 1600
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI070DH03QP1.
Location/Qualifiers
1. 1112
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="CSODI070YP06"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 414 a 149 c 189 g 262 t 98 others
ORIGIN
Query Match 19.2%; Score 45.8; DB 9; Length 1112;
Best Local Similarity 38.0%; Pred. No. 0.78;
Matches 90; Conservative 37; Mismatches 110; Indels 0; Gaps 0;
QY 2 ATTGGATACGTTGGAAGGATATTTCGTTGGGAATGGGATGAAGATACGCTAGATG 61
DB 835 MTTTTTTTTTTGTAATAAAWAKAAWTTTAAWTKWTATTTTTAAWAGAAAGAAARRGGK 894
QY 62 AGTTGGGAAAAAGTGTAGAACATATCGGTGAAGATGAACCGTTATTGTTGAAAAATC 121
DB 895 TTTTAAAAAARRKKGAARAAATTTTTTTTAKATTTTGTGKAAAAAATWAA 954
QY 122 TACTGTTAACTTCAATCAGGAAAAACATGAAGAGTTAGGAAGTTGATTTATCAGTTGA 181
DB 955 ATTAGGKAKAWGTTTATGCAAAAAAAGAGAWTTTGAATKAWWATTTTWTWT 1014
QY 182 TAGAGTTAGATTTCTCGAACTTTTGTGAGGATTTCTATGAAAACTGGAAGCAATT 238
DB 1015 TTAAWAAATTTTIDWWWRGGTTTTTATAAAAKGGGTTTTAAAAAATAAAW 1071
RESULT 3
BX361014 989 bp mRNA linear EST 05-MAY-2003
LOCUS
DEFINITION
BX361014 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI078YP09 5-PRIME, mRNA sequence.
ACCESSION
BX361014.1 GI:30382486
VERSION
EUT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 989)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10967.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI078CH05QP1&cluster=10967.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI078CH05QP1.
Location/Qualifiers
1. 989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI078YP09"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 358 a 127 c 176 g 152 t 176 others
ORIGIN

```

Query Match      18.6%; Score 44.2; DB 13; Length 989;
Best Local Similarity 32.9%; Pred. No. 1.9;
Matches 69; Conservative 52; Mismatches 89; Indels 0; Gaps 0;

QY 4 TTGGATAGCTTGGAAAGGATATTGTTGGGAATGGATGAAGATGAAGCGTAGATGAG 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
686 TRGGGARGKGGKTDAAWAAAWTWTTTKGTGABAAAAAARARAGGGKGTDWAAAA 745
QY 64 TTGGGAAAAAAGTGTAGACATATCGTGAAGATGAACCGTATTGTTGAAAAATCTA 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
746 WTGTGAAAAAAAKTKGGRNRATATAARGGRKWKTKWKTADAADDDKTTAAKATADA 805
QY 124 CTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTATCAGTTGATA 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
806 ADATTTTXXTKAAAAAAGGKCAAAWKGGRGAKKATGKTKAAAAGWTTTXXKTT 865
QY 184 GAGTTAGATTTTCTGGAACCTTTGTGAGGG 213
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
866 TWTWRAAATTTWKAIAAARTTKRAGARR 895

RESULT 4
LOCUS BX369737
DEFINITION BX369737 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC006YG10 5-PRIME, mRNA sequence.
ACCESSION BX369737
VERSION BX369737.1 GI:30453837
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 905)
JOURNAL Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
COMMENT Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3358.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cwuster.cgi?seq=CS0BAG0052A04_CS00409_1.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Paradise Avenue Genoscope sequence ID : CS0BAG0052A04_CS00409_1.
FEATURES
source
1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC006YG10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 365 a 119 c 163 g 251 t 7 others
ORIGIN
18.2%; Score 43.2; DB 13; Length 905;
Best Local Similarity 55.9%; Pred. No. 3.2;
Matches 104; Conservative 0; Mismatches 78; Indels 4; Gaps 1;

QY 27 TTGTATGGGAATGGATGAAGATGAACCGTAGATGAGTTGGGAAAAAAGATGTTAGACA 86
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
381 TTTAATTTGGAAATTTTGGAAAGAACCTTAATTTTGGGGGGGGGAAAAAATTGAAAAAC 440

Query Match      18.2%; Score 43.2; DB 13; Length 905;
Best Local Similarity 55.9%; Pred. No. 3.2;
Matches 104; Conservative 0; Mismatches 78; Indels 4; Gaps 1;

QY 27 TTGTATGGGAATGGATGAAGATGAACCGTAGATGAGTTGGGAAAAAAGATGTTAGACA 86
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
381 TTTAATTTGGAAATTTTGGAAAGAACCTTAATTTTGGGGGGGGGAAAAAATTGAAAAAC 440

```

```

QY 87 TATCGGTAGAAATGAACCGTATTGTTGAAAAATCTACTGGTTAACTTCAATCAGGAAA 146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
441 TACTAAAAAAATTTAAAGGTTTAAAGGAAAAAATAATTTTAAAGGAAAAAAGGTAA 500
QY 147 ACA----TGAAAGATAGTAGAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGAAC 202
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
501 AAAAAAGTGAAGAAGCTTGCTGTTGAAAAAGTTTGTCTGGGTTGATTTATGAAAAT 560
QY 203 TTTTGT 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
561 TTTTIT 566

RESULT 5
LOCUS AQ222887/c
DEFINITION HS 2013_A1_B04_T7 CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2013 Col=7 Row=C, genomic survey
sequence.
ACCESSION AQ222887
VERSION AQ222887.1 GI:3636500
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 353)
JOURNAL Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
COMMENT Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2013 row: C column: 7
Class: BAC ends
High quality sequence stop: 353.
FEATURES
source
1..353
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2013 Col=7 Row=C"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 92 a 94 c 46 g 121 t
ORIGIN
Query Match      18.1%; Score 43; DB 28; Length 353;
Best Local Similarity 53.2%; Pred. No. 2.7;
Matches 91; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 AAAAGGGATATTGATTGGGAATGGCATGAAGATGAGTTAGATGAGTTGGGAAAAAAG 76
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 AAAGTAAAAAGTTATATCAGGTTTGAAGGAGACAACTACTAAATGTTAGGAGAAAAG 112
QY 77 TCTTAGACATATCGTGAAGATGAACCGTTATTCTTGGAAAAATCTACTGTTACTTCA 136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
111 TCAATAAATTTTGGTGGACATCCATCATCAGCGTTAGTAAATATTACTTTTAAATCG 52
QY 137 ATCAGGAAAAACATCAAGAAGTTAGGAAGTTGATTATCAGTTGATAGT 187

```

```
Db 51 GTAGAGGAACCTCAATCAGCAGCAATAGCTCAATTTTATAAGATGATGATT 1

RESULT 6
AQ506302
LOCUS
DEFINITION
PCPI-11-278E6.TU RPCI-11 Homo sapiens genomic clone RPCI-11-278E6,
genomic survey sequence.
ACCESSION
AQ506302
VERSION
AQ506302.1 GI:4711049
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 575)
Zhaoc,S., Adams,W.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished
Other GSSs: RPCI-11-278E6.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hunger/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
1..575
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/db_xref="GDB:7606469"
/db_xref="taxon:9606"
/clone="RPCI-11-278E6"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 242 a 99 c 99 g 135 t
ORIGIN
Query Match 18.1%; Score 43; DB 28; Length 575;
Best Local Similarity 51.3%; Pred. No. 3.1;
Matches 100; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 43 ATGAGATAAGCGTAGATGAGTTGGGAAAAAAGTTGTAGAACATATCGGTAAAGATCAA 102
|||||
Db 183 ATGATTTATAGCTAAGTATTTTCGACAGAGTACCAGAACATTATTGGAGAGAGAA 242
QY 103 CCGTTATTGTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATGAAGAAGTTAGG 162
Db 243 CAGTCTCTTAAATAATGTCACCGGAAAAATTTGAATTCATATACAGAAAGAGAAACT 302
QY 163 AGTTGCAITTCAGTTAGATGATTTCTCGGAACCTTTTGTGAGGATTCCTATGA 222
Db 303 AGATTACTCTCTCTCACTATATACAAAAATTAACATCAAGTGGCTTAAGACTTACATGT 362
QY 223 AAAACTGGAAGCAAT 237
Db 363 AAGACCTGAAACTAT 377
```

```
RESULT 7
CNS0006J
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR01M22 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL062049
VERSION
AL062049.1 GI:4938511
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Direct Submission
Genoscope.
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR01M22"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 333 a 162 c 148 g 177 t 281 others
ORIGIN
Query Match 18.1%; Score 43; DB 29; Length 1101;
Best Local Similarity 24.7%; Pred. No. 3.8; Indels 0; Gaps 0;
Matches 55; Conservative 79; Mismatches 89; Indels 0; Gaps 0;
QY 15 GGAAAAGGGATATTGATTGGGATGGGATGAAGATAGCGGTAGATGAGTTGGGAAAAA 74
|||||
Db 727 GRANRAAAAKTKKKTKTKRRAAGRRARRAWAGAAAAAARAADAATKTGTWAKAA 786
|||||
QY 75 AGTGTAGAACATATCGGTAAAGTAAGACCGTTATGTTGAAAAATCTACTGTTAATT 134
|||||
Db 787 WRTAAWKDWKATDAAKAAKAAARTDWTAKAKTKWGAATGAWGAGAAAGRRWGK 846
|||||
QY 135 CAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCATGTTGATAGATTAGATT 194
|||||
Db 847 KGDGTPARPARAGAGDGDWDKAAWAAWAAATATSWDTATADDDRRKAWDKRAAADR 906
|||||
QY 195 TTGGAACCTTTTGTGAGGATTCCTATGAAAAAACTGGAAGCAAT 237
|||||
Db 907 DXDKGRAGGTCGWRRRAPARTKAKAKERGAGAKAKKRAK 949
|||||
RESULT 8
B53839
LOCUS
DEFINITION
CIT-HSP-2013C7.TF CIT-HSP Homo sapiens genomic clone 2013C7,
genomic survey sequence.
ACCESSION
B53839
```


VERSION	B53839.1	GI:2608173
KEYWORDS	GSS.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. & Venter,J.C.	
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building	
JOURNAL	Unpublished	
COMMENT	Other GSSs: CIT-HSP-2013C7.TR Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html Seq primer: M13-21 Class: BAC ends.	
FEATURES	Location/Qualifiers	
source	1..483	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="GBS:7042233"	
	/db_xref="taxon:9606"	
	/clone="2013C7"	
	/sex="Male"	
	/cell_type="Sperm"	
	/clone_lib="CIT-HSP"	
	/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"	
BASE COUNT	142 a 111 c 75 g 155 t	
ORIGIN		
Query Match	17.8%; Score 42.4; DB 28; Length 483;	
Best Local Similarity	52.9%; Pred No. 4.1;	
Matches	91; Conservative 0; Mismatches 81; Indels 0; Gaps 0;	
QY	17 AAAGCGATATTGATTGGGAATGGATGAAGATAGCGTAGAGTGGGAAAAAAG 76	
Db	268 AAAGTAAAAAGTTATATACAGGTGGAGGAAACAACTACTAAATGTTAGGAGAAAAG 209	
QY	77 TGTAGACATATCGTAGATGAACCTGTTATGTTGAAAAAATCTAGTGTAACTTCA 136	
Db	208 TCAATATATTTTGGTGGCAATCCATCATACGCGTTAGTAAATATACATTTTAATCG 149	
QY	137 ATCAGGAAACATGAAGAAGTTAGGAAGTTGATTATACGTTGATAGAGTT 188	
Db	148 GTAGAGGAACCTCAAATCACACCAATAGTCAATTTTAAGATGATGTTT 97	
RESULT 9		
CB285716		
LOCUS	CMD22_H07_61 UMMNPM3 Sus scrofa cDNA clone PPSUBLIB_06H07 5', mRNA	
DEFINITION	sequence.	
ACCESSION	CB285716	
VERSION	CB285716.1	
KEYWORDS	EST.	
SOURCE	Sus scrofa (pig)	
ORGANISM	Sus scrofa	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
AUTHORS	Dvorak,C.M.T., Hyland,K.A., Zhang,Y., Fahrenkrug,S.C. and Murtaugh,M.P.	

FEATURES

source		1. 721		http://www.genoscope.cns.fr/	
		/organism="Sus scrofa"		cgi-bin/cluster.cgi?seq=CSIAF002H04NP1&cluster=8830.r. Contact :	
		/mol_type="mRNA"		Feng Liang Email : fliang@lifetech.com URL :	
		/db_xref="taxon:9823"		http://fulllength.invitrogen.com/ Invitrogen Corporation 1600	
		/clone="PPSUBLIB_06H07"		Paraday Avenue Genoscope sequence ID : CSIAF002H04NP1.	
		/tissue_type="Peyer's patch"		Location/Qualifiers	
		/clone_lib="UNMPPM3"		1. 1175	
		/note="Organ: small intestine, jejunum; Vector: pCMVSPORT6		/organism="Homo sapiens"	
		; Site 1: NotI; Site 2: SalI; Jejunal Peyer's patches were		/mol_type="mRNA"	
		isolated from healthy, 4-6 week old cross-bred pigs. RNA		/db_xref="taxon:9606"	
		was extracted either immediately or after 3 hours		/clone="CSODF008YB06"	
		stimulation in an Ussing chamber with one of the following		/tissue_type="FETAL BRAIN"	
		treatments: 1) no treatment, 2) Salmonella choleraesuis		/dev_stages="fetal"	
		vaccine strain SC-54, 3) phorbol myristate acetate,		/clone_lib="Homo sapiens FETAL BRAIN"	
		canavanin A, and 8-bromo-cyclic AMP, and 4)		/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA	
		lipopolysaccharide and cholera toxin. Each treatment was		was primed with a NotI-oligo(dT) primer. Five prime end	
		performed in the presence and absence of cycloheximide.		enriched, double-strand cDNA was digested with NotI and	
		Purified poly A + RNA from each of the treated tissues		cloned into the Not I and EcoRV sites of the pCMVSPORT 6	
		(2-4) was combined together, reverse transcribed, and		vector. Library was not normalized."	
		cloned in to pCMVSPORT6 to make a library of approximately		BASE COUNT	
		530,000 recombinant clones with an average insert size of		410 a 121 c 115 g 357 t 172 others	
		1.0 kbp. Poly A+ RNA from freshly isolated and non-treated		ORIGIN	
		tissue from an Ussing chamber (treatment 1) was cloned in		Query Match	
		the same manner to produce an unstimulated cDNA library of		17.6%; Score 41.8; DB 13; Length 1175;	
		approximately 900,000 clones with an average insert size		Best Local Similarity	
		of 1.5 kbp. Equal portions of the two libraries were		35.5%; Pred. No. 7.5; Mismatches 93; Indels 0; Gaps 0;	
		pooled and then subtracted with porcine ST fibroblast RNA		Matches 75; Conservative 43;	
		to create a subtracted porcine Peyer's patch library of		QY 12 GTTGGAAAAGGATATTTCATTTGGCAATGGGATGAAGATAGCGTAGATCAGTTGGGAA 71	
		approximately 6,000 clones with an average insert size of		Db 813 VKVGNVTRAGAGATGGATRAARRAATTCATTGATATGGATBDTTRATTTGRRW 872	
		1.0 kbp."		QY 72 AAAAGTGTAGAACATATCGGTAGACATGAACCGTTATTGTTGAAAATCTACTGGTTAA 131	
		BASE COUNT		Db 873 KTTTTRTTTTRATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 932	
		227 a 120 c 124 g 248 t 2 others		QY 132 CTTCAATCAGGAAAACATGAAGATTAGGAAGTTGATTATTCAGTTGATAGATTAGA 191	
		ORIGIN		Db 933 TWWRTTWTATKSNARATTTWAAAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 992	
		Query Match		QY 192 TTTTCGGAACTTTTGTGAGGATTCCTATGA 222	
		17.6%; Score 42; DB 14; Length 721;		Db 993 TTTTTTTATATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1023	
		Best Local Similarity		RESULT 11	
		55.0%; Pred. No. 5.8; 50; Indels 0; Gaps 0;		AZ549439/c	
		Matches 72; Conservative 0; Mismatches 50; Indels 0; Gaps 0;		LOCUS	
		QY 79 TTAGAACAATATCGGTAAAGATGAACCGTTATTGTTGAAAATCTACTGGTTAACTCAAT 138		DEFINITION	
		Db 218 TAAAAACATCTCGGTCAAGATCTGCACCTTTTGTATCTTCAGGTTCTCGTTAGTAAACT 277		ENTDG77F Entamoeba histolytica Sheared DNA Entamoeba histolytica	
		QY 139 CAGGAAAACATGAGAGATTAGGATTCATTATTCAGTTGATAGATTAGATTCTTG 198		Genomic, genomic survey sequence.	
		Db 278 CAAGGAGAAATAAAAAAGTACGTAATGAATTAATTTTTCATTTGGTTAAATTTTATG 337		ACCESSION	
		QY 199 GA 200		AZ549439	
		Db 338 AA 339		VERSION	
		RESULT 10		AZ549439.1 GI:11174048	
		BX423299		GSS	
		BX423299 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone		Entamoeba histolytica	
		CSODF008YB06 3-PRIME, mRNA sequence.		Entamoeba histolytica	
		DEFINITION		Eukaryota; Entamoebidae; Entamoeba.	
		ACCESSION		1 (bases 1 to 862)	
		BX423299		Loftus,B., Van Aken,S. and Fraser,C.	
		BX423299.1 GI:30655486		Determination of clone end sequences from Entamoeba histolytica	
		EST.		HM1:IMSS sheared DNA library	
		SOURCE		Unpublished	
		Homo sapiens (human)		Contact: Brendan J Loftus	
		Homo sapiens		Department of Eukaryotic Genomics	
		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;		The Institute for Genomic Research	
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		9712 Medical Center Dr., Rockville, MD 20850, USA	
		1 (bases 1 to 1175)		Tel: 301 838 0208	
		Li W.B., Gruber C., Jessee, J. and Polaves, D.		Fax: 301 838 3543	
		Full-length cDNA libraries and normalization		Email: b.loftus@igr.org	
		Unpublished		Clones are derived from the Entamoeba histolytica HM1:IMSS sheared	
		Contact: Genoscope		DNA library	
		Genoscope - Centre National de Sequencage		Seq primer: M13-Forward	
		BP 191 91006 EVRY cedex - France		Class: shotgun	
		Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		High quality sequence start: 23	
		Library was constructed by Life Technologies, a division of		High quality sequence stop: 808.	
		Invitrogen. This sequence belongs to sequence cluster 8830.r For		Location/Qualifiers	
		more information about this cluster, see		FEATURES	

```

source
1..862
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pROSL; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
BASE COUNT      197 a 156 c 98 g 411 t
ORIGIN

Query Match      17.5%; Score 41.6; DB 28; Length 862;
Best Local Similarity 57.8%; Pred. No. 7.7;
Matches 74; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 37 AATGGATGAGATAGCGTACATGAGTGGGAAAAGTGTAGACATATCGGTAAG 96
|||||
DB 467 AATGAAAGAAAAGAGTAGTGAAGAAATAGAGGAAAAAATATAGAGATATGAAGAG 408
|||||

QY 97 AATGAACCGTATTGTTGAAAATCTACTGTTTCACTTCAATCGAGGAAAACATGAAGAA 156
|||||
DB 407 AAGAAGAAATGAAAGAAAAGAACTATTGAATTAATAAATAAACCAGAAATAA 348
|||||

QY 157 GTTAGGAA 164
|||||
DB 347 TTAATGAA 340

RESULT 12
CNS039G/c.
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR0810 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL063921
VERSION
AL063921.1 GI:4941778
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyceta; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
REFERENCE
Direct Submission
AUTHORS
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammos in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be

```

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR0810"
/clone_lib="RPCI-98"
/notes="end : TET3"
BASE COUNT      201 a 64 c 131 g 202 t 503 others
ORIGIN

Query Match      17.5%; Score 41.6; DB 29; Length 1101;
Best Local Similarity 14.5%; Pred. No. 8.2;
Matches 34; Conservative 128; Mismatches 72; Indels 0; Gaps 0;

QY 2 ATTGGATAGTTGGAAAAGGATATTGATTGGGAATGGATGAAGAATAAGCGTAGATG 61
|||||
DB 970 DKDKRGDDDDKGGKKKAAKAAKWATKWDWDDWDDKWDGAKDRKADDDGAGDKD 911
|||||

QY 62 AGTTGGGAAAAGTGTTAGACATATCGTAAAGTACCGTTATTGTTGAAAAATC 121
|||||
DB 910 DDKGKXADDDDTGTDKDDDDKDKWDDWAKGTWGDATWAAATDWWWGWADADWWTWD 851
|||||

QY 122 TACTGTTTAACCTCAATCAGGAAAACATGAAAGTTAGGAAGTTGATTATCAGTTGA 181
|||||
DB 850 AAADDDWADDWDWAWKNDWAWGARTADREDWDGAKRGARKRDRKADKRD 791
|||||

QY 182 TAGAGTTAGATTCTTGGAACTTTTGTGAGGATTTCTATGAAAACCTGAAGCA 235
|||||
DB 790 ADDRDDAATTTTTRDTTDDKWKKTDTWTRWAADRTWDRDDDDDRDRAGTA 737
|||||

RESULT 13
AV885336/c
LOCUS
DEFINITION
AV885336 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone rcitb19114 3', mRNA sequence.
ACCESSION
AV885336
VERSION
AV885336.1 GI:16872860
KEYWORDS
EST.
SOURCE
Ciona intestinalis
ORGANISM
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cnidaria; Ciona.
1 (bases 1 to 641)
REFERENCE
Satoh N., Satou Y., Kohara Y. and Shin-i, T.
Expressed genes in Ciona intestinalis
Unpublished
CONTACT: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..641
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcitb19114"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud
embryo"
BASE COUNT      210 a 123 c 107 g 200 t 1 others
ORIGIN

Query Match      17.3%; Score 41.2; DB 9; Length 641;
Best Local Similarity 53.0%; Pred. No. 8.8;
Matches 88; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```

QY 42 GATGAAGTAAAGCGTAGATGAGTTGGGGAAAAAGTTGTTAGAACATATCGGTAAAGATGA 101
 Db 605 GCTCAAGTTCTGGTAAAAAAGGTTGTCAATAGGAATTTGCACCTTTGGCCAGAACGGT 546
 QY 102 ACCGTTATTGTTGAAAAATCTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAG 161
 Db 545 TGTGTTTTTGTGAAAAATCTACAAATAAGTTGTTGATGTTGAGCCCGCCAGCAGTAAG 486
 QY 162 GAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGAATTTTG 207
 Db 485 CCTGTTGATACACAGGTGCTGGAGATGCATTTACTGGAGCTCTTG 440

RESULT 14
 AV849915/c
 LOCUS
 DEFINITION AV849915 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone rcic104e15 3', mRNA sequence.
 ACCESSION AV849915
 VERSION AV849915.1 GI:16832631
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

REFERENCE
 AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
 source
 1..679
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcic104e15"
 /tissue_type="whole animal"
 /dev_stage="cleavage stage embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, cleavage stage embryo"
 BASE COUNT 219 a 126 c 109 g 220 t 5 others
 ORIGIN

Query Match 17.3%; Score 41.2; DB 9; Length 679;
 Best Local Similarity 53.0%; Pred.No.9;
 Matches 88; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 QY 42 GATGAAGTAAAGCGTAGATGAGTTGGGGAAAAAGTTGTTAGAACATATCGGTAAAGATGA 101
 Db 641 GCTCAAGTTCTGGTAAAAAAGGTTGTCAATAGGAATTTGCACCTTTGGCCAGAACGGT 582
 QY 102 ACCGTTATTGTTGAAAAATCTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAG 161
 Db 581 TGTGTTTTTGTGAAAAATCTACAAATAAGTTGTTGATGTTGAGCCCGCCAGCAGTAAG 522
 QY 162 GAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGAATTTTG 207
 Db 521 CCTGTTGATACACAGGTGCTGGAGATGCATTTACTGGAGCTCTTG 476

RESULT 15
 BW092427/c
 LOCUS
 DEFINITION BW092427 Nori Satoh unpublished cDNA library, young adult Ciona
 intestinalis cDNA clone rcic103p01 3', mRNA sequence.
 ACCESSION BW092427
 VERSION BW092427.1 GI:24268229

KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.
 REFERENCE
 AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
 TITLE Expressed genes in Ciona intestinalis (2002c)
 JOURNAL Unpublished
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
 source
 1..737
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcic103p01"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young adult"
 BASE COUNT 236 a 148 c 123 g 230 t
 ORIGIN

Query Match 17.3%; Score 41.2; DB 13; Length 737;
 Best Local Similarity 53.0%; Pred.No.9.2;
 Matches 88; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 QY 42 GATGAAGTAAAGCGTAGATGAGTTGGGGAAAAAGTTGTTAGAACATATCGGTAAAGATGA 101
 Db 617 GCTCAAGTTCTGGTAAAAAAGGTTGTCAATAGGAATTTGCACCTTTGGCCAGAACGGT 558
 QY 102 ACCGTTATTGTTGAAAAATCTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAG 161
 Db 557 TGTGTTTTTGTGAAAAATCTACAAATAAGTTGTTGATGTTCCAGCCCGCCAGCAGTAAG 498
 QY 162 GAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGAATTTTG 207
 Db 497 CCTGTTGATACACAGGTGCTGGAGATGCATTTACTGGAGCTCTTG 452

Search completed: November 15, 2003, 08:02:50
 Job time : 1590.72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:06:39 ; Search time 37.0704 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-96

Perfect score: 238
Sequence: 1 AATTGGATACCTTGGAAAA.....ATGAAAACTGGAAGCAATT 238

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220591566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/FCUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41.8	17.6	7218	1	US-08-232-463-14
C 2	40	16.8	1564976	4	US-08-916-421B-1
C 3	38.2	16.1	1287	4	US-09-107-532A-3495
C 4	35.8	15.0	1581	4	US-09-412-778-1
C 5	35.8	15.0	1058	4	US-08-213-419B-20
C 6	35.8	15.0	3107	4	US-08-213-419B-1
C 7	35.8	15.0	6124	4	US-08-213-419B-3
C 8	34.8	14.6	1194	4	US-09-134-001C-2667
C 9	34.6	14.5	3450	4	US-09-904-615-44
C 10	34.4	14.5	1032	4	US-09-107-532A-2955
C 11	34.2	14.4	1589	3	US-08-747-574-2
C 12	34	14.3	1068	3	US-08-913-842-12
C 13	33	13.9	2981	1	US-08-257-073-2
C 14	33	13.9	2981	2	US-08-184-009-119
C 15	33	13.9	2981	2	US-08-458-356-119
C 16	33	13.9	2981	3	US-08-460-736-119
C 17	33	13.9	2981	4	US-09-535-370-119
C 18	32.8	13.8	3857	3	US-09-347-114A-80
C 19	32.8	13.7	3001	4	US-09-539-333D-229
C 20	32.6	13.7	3001	4	US-09-328-352-2460
C 21	32.2	13.5	2007	4	US-09-328-352-2460
C 22	32	13.4	570	4	US-09-134-001C-964
C 23	32	13.4	1287	4	US-09-601-198-17
C 24	32	13.4	1641	4	US-08-583-276-18
C 25	32	13.4	4659	2	US-08-961-527-254
C 26	31.8	13.4	2789	4	US-08-961-527-254
C 27	31.8	13.4	1664976	4	US-08-916-421B-1

28	31.6	13.3	1485	1	US-08-468-036-8	Sequence 8, Appl
29	31.6	13.3	1485	2	US-08-378-843-8	Sequence 8, Appl
30	31.6	13.3	3098	1	US-08-447-500-1	Sequence 1, Appl
31	31.6	13.3	3098	1	US-08-454-097-1	Sequence 1, Appl
32	31.6	13.3	3098	1	US-08-447-408-1	Sequence 1, Appl
33	31.6	13.3	3098	1	US-08-453-866-1	Sequence 1, Appl
34	31.6	13.3	3098	3	US-08-185-359-1	Sequence 1, Appl
35	31.6	13.3	4061	4	US-09-620-312D-363	Sequence 363, App
C 36	31.4	13.2	1830121	4	US-09-557-884-1	Sequence 1, Appl
C 37	31.4	13.2	1830121	4	US-09-643-990A-1	Sequence 1, Appl
38	31	13.0	630	4	US-09-328-352-3314	Sequence 3314, Ap
39	31	13.0	1720	4	US-09-227-357-139	Sequence 139, App
40	31	13.0	2095	4	US-09-227-357-31	Sequence 31, Appl
C 41	31	13.0	2726	1	US-08-461-823-1	Sequence 1, Appl
C 42	31	13.0	3480	1	US-08-920-812-11	Sequence 11, Appl
C 43	31	13.0	3480	1	US-08-920-827-11	Sequence 11, Appl
C 44	31	13.0	3480	1	US-08-921-177-11	Sequence 11, Appl
C 45	31	13.0	3480	1	US-08-362-577C-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Folsy & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-F1s
; US-08-232-463-14.

Query Match 17.6%; Score 41.8; DB 1; Length 7218;

```
Best Local Similarity 3.08; Pred. No. 0.02;
Matches 7; Conservative 143; Mismatches 85; Indels 0; Gaps 0;

QY 2 ATTGGATCGTTGGAAGGATATTTGATTGGGAATGGGATGAAGATACCGTAGATG 61
    ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1444 AATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1385
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 AGTTGGGAAAAGTGTAGACATATCGGTAAGATCAACCGTTATTGTGAAAATC 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1384 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1325
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 TACTGTTAACTTCAATCAGGAAACATGAAGAAGTTAGGAAGTTATTTATCAGTTGA 181
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1324 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1265
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 TACAGTAGATTTCTGGACITTTGTGAGGATCTATGAATACTGAAGCA 236
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1264 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1210
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2822)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
```

NAME/KEY: misc feature
LOCATION: (779576)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1113981)..(1113981)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1563020)..(1563020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g

US-08-916-421B-1

Query Match 16.8%; Score 40; DB 4; Length 1664976;
Best Local Similarity 55.9%; Pred. No. 0.22;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 70 AAAAAGTGTAGAACATATCGGTAGATGAACCGTTATTGTTGCAAAATCTACTGTT 129
DB 1095978 AACTAAATTTTGCATACCTCGGGAGAGATAGCAAAATAGTGGAAATTTATTGGTC 1096037
QY 130 AACTTCAATCAGGGAACATGAAGAGTTAGGAAGTTGATTATCAGTTGATAGAGTTA 189
DB 1096038 TATTCCTAAGTGAAGAAATACCGCAAAATTTGGATTATTAATGAACATTAGATAGAGTTA 1096097
QY 190 GATTTTCTGGAACTTT 205
DB 1096098 CAGACGAGGAGCTTT 1096113

RESULT 3
US-09-107-532A-3495
; Sequence 3495, Application US/09107532A

Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3495:
SEQUENCE CHARACTERISTICS:
LENGTH: 1287 base pairs
TYPE: nucleic acid
STRANDEDNESS: Double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1287
SEQUENCE DESCRIPTION: SEQ ID NO: 3495:
US-09-107-532A-3495

Query Match 16.1%; Score 38.2; DB 4; Length 1287;
Best Local Similarity 56.9%; Pred. No. 0.12;
Matches 70; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 78 GTTAGACATATCGGTAAAGATGAACCGTTATTGTTGAAAAATCTACTGTTAACTCAA 137
DB 558 GTTGAATAATATCGTAAAAAAGACTTCTTAATCTTGAATAATCTCTTAGGAAATTTGAA 617
QY 138 TCAGGAAACATGAAGAGTTAGGAAGTTGATTATCAGTTGATAGAGTTAGATTCTTCT 197
DB 618 GTACAGCAGCTATTAAGAGTAGTAAGTAATTTTGTCAATACGATTTTTCAGAAATTTTC 677
QY 198 GGA 200
DB 678 GGA 680

RESULT 4
US-09-412-779-1
; Sequence 1, Application US/09412779
; Patent No. 6352843
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena

APPLICANT: Warren, Richard L.
TITLE OF INVENTION: Vsx
FILE REFERENCE: GM10244
CURRENT APPLICATION NUMBER: US/09/412,779
CURRENT FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 591
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-09-412-779-1

Query Match 15.0%; Score 35.8; DB 4; Length 591;
Best Local Similarity 50.3%; Pred. No. 0.45;
Matches 88; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 22 GGATATTTGGAATGGGATGAAGATAGCGTAGATGTTGGGAAAAAAGTGTTA 81
DB 238 GGGTATGATATGCTAAAGTAGTAAGTAAACACACACGTCGAAAAATTTGGAAAAATGATTCAG 297
QY 82 GAACATATCGGTGAAGATGAACCGTTATTGTGAAAAATCTACTGTTAACTTCAATCAG 141
DB 238 GAATATATACTAGACAGAGCAATTTGCAATTAGTTATTCAATTAGTTGATTTAAGACAT 357
QY 142 GGAACATGAAGAGTTAGGAAGTTGATTTATCAGTTGATAGAGTTAGATTTTC 196
DB 358 GATCCACACAGATGATCTCTTAATGTACAATTATTGAAACATTTTGATATTC 412

RESULT 5
US-08-213-419B-20
Sequence 20, Application US/08213419B
Patent No. 6333406
GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: JII-002CNC
CURRENT APPLICATION NUMBER: US/08/213,419B
CURRENT FILING DATE: 1994-03-14
PRIOR APPLICATION NUMBER: US 07/870,506
PRIOR FILING DATE: 1992-04-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 1058
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-08-213-419B-20

Query Match 15.0%; Score 35.8; DB 4; Length 1058;
Best Local Similarity 59.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 98 ATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATCAAGAAG 157
DB 20 ATTTACAAAAATGTTATATAATGTACAGGAGAAAGTCAACAGGTAATACAGGAGGAGG 79
QY 158 TTAGGAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGA 200
DB 80 TCAAGCAGGTAATACAGTAGGAGATCAAGCAGGTAATACAGGAGGAGG 122

RESULT 6
US-08-213-419B-1
Sequence 1, Application US/08213419B
Patent No. 6333406
GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: JII-002CNC

CURRENT APPLICATION NUMBER: US/08/213,419B
CURRENT FILING DATE: 1994-03-14
PRIOR APPLICATION NUMBER: US 07/870,506
PRIOR FILING DATE: 1992-04-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3107
TYPE: DNA
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: CDS
LOCATION: (104)..(3070)
US-08-213-419B-1

Query Match 15.0%; Score 35.8; DB 4; Length 3107;
Best Local Similarity 59.2%; Pred. No. 0.67;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 98 ATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATCAAGAAG 157
DB 145 ATTTACAAAAATGTTATATAATGTACAGGAGAAAGTCAACAGGTAATACAGGAGGAGG 204
QY 158 TTAGGAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGA 200
DB 205 TCAAGCAGGTAATACAGTAGGAGATCAACAGGTAATACAGGAGGAGG 247

RESULT 7
US-08-213-419B-3
Sequence 3, Application US/08213419B
Patent No. 6333406
GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: JII-002CNC
CURRENT APPLICATION NUMBER: US/08/213,419B
CURRENT FILING DATE: 1994-03-14
PRIOR APPLICATION NUMBER: US 07/870,506
PRIOR FILING DATE: 1992-04-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 6124
TYPE: DNA
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: CDS
LOCATION: (2407)..(2439)
NAME/KEY: CDS
LOCATION: (2598)..(3404)
NAME/KEY: CDS
LOCATION: (3580)..(3720)
NAME/KEY: CDS
LOCATION: (3850)..(5835)
US-08-213-419B-3

Query Match 15.0%; Score 35.8; DB 4; Length 6124;
Best Local Similarity 59.2%; Pred. No. 0.79;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 98 ATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATCAAGAAG 157
DB 2606 ATTTACAAAAATGTTATATAATGTACAGGAGAAAGTCAACAGGTAATACAGGAGGAGG 2665
QY 158 TTAGGAAGTTGATTTATCAGTTGATAGAGTTAGATTTTCTGGA 200
DB 2666 TCAAGCAGGTAATACAGTAGGAGATCAACAGGTAATACAGGAGGAGG 2708

RESULT 8
US-09-134-001C-2667

; Sequence 2667, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2667
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2667

Query Match 14.6%; Score 34.8; DB 4; Length 1194;
Best Local Similarity 50.0%; Pred. No. 0.99; Mismatches 0; Gaps 0;
Matches 87; Conservative 0; Indels 87

QY 10 ACGTTGGAAGGATATTGTTGGGATGGGATGAAGATAGCGTAGATGAGTTGGG 69
Db 179 ACTTAGTAACTTAAAGAGAGAGTGATAAAGAGGACTTTCACCTTAAGCCGTGCTG 238
QY 70 AAAAAGCTTTAGACATATCGTGAAGATGACCGTTATTGTTGAAAAAATCTACTGTT 129
Db 239 AAAAATTATCTAAGAAATTAGTAAAGAGATTATTGCTTCCTGAAACACGTCGCGAA 298
QY 130 AACTTCAATCAGGGAACATGAAGAAGTTAGGAAGTTGATTATTCAGTTGATA 183
Db 299 AACTCGAACACGAATTGAACCTTAATGAAGGCACTTTATTAGTTGAA 352

RESULT 9
US-09-904-615-44/c
; Sequence 44, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 44
; LENGTH: 3460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-44

Query Match 14.5%; Score 34.6; DB 4; Length 3460;
Best Local Similarity 49.7%; Pred. No. 1.4; Mismatches 89; Indels 89; Gaps 0;
Matches 88; Conservative 0

QY 1 AATTGGATACGTTGGAAAGGATATTGTTGGGAATGGGATGAAGATAGCGTAGAT 60
Db 2072 AATATACAAACACTAGGTAAATTCAAATTTCTTTGTAATGCGATGGAAGTCTACTGAGAA 2013
QY 61 GAGTTGGGAAAAAGTGTGAGACATATCGTGAAGATGACCGCTTATTGTTGAAAAAT 120
Db 2012 GCAGAAAAAGAACACCATTTTAAACATTTATCTGGCTAGTCAACTGGTAGTGGGTATAAAT 1953

QY 121 CTACTGTTAACTTCAATCAGGGAAAAACATGAAGAAGTTAGGAAGTTGATTATCAG 177
Db 1952 GTACTGTAAGGTCAGGCCACACCTAACCTCAATATGAGCCAATTAGTTTCACAG 1896

RESULT 10
US-09-107-532A-2955
; Sequence 2955, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2955:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1092
; SEQUENCE DESCRIPTION: SEQ ID NO: 2955:
US-09-107-532A-2955

Query Match 14.5%; Score 34.4; DB 4; Length 1092;
Best Local Similarity 51.3%; Pred. No. 1.2; Mismatches 0; Gaps 0;
Matches 80; Conservative 0; Indels 76

QY 31 ATTGGGAATGGGATGAAGATAGCGTAGATGTTGGGAAAAAAGTTAGAACATATC 90
Db 737 ATTGTGTCATAATGCACACAGCGTGGAGAGATCTAGACAAAAATGATTTAGCAGCTA 796
QY 91 GGTAAAGATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCATCAGGGAACAT 150
Db 797 GCTTCAAGCAAGTGTGATGTTGTTATTGATTACTAAAAACACITCGAGCTTGTCAAAAT 856
QY 151 GAAGAAGTTAGGAAGTTGATTTATCATTCAGTTGATAGAG 186
Db 857 ATCCAGTCAAGCAACTGGTTGTAGCTGGAGGTAG 892

RESULT 11

US-08-747-574-2
; Sequence 2, Application US/08747574
; Patent No. 6015939
; GENERAL INFORMATION:
; APPLICANT: CALGENE, INC.
; TITLE OF INVENTION: PLANT VDE GENES AND
; TITLE OF INVENTION: METHODS RELATED THEREO
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,574
; FILING DATE: No. 6015939ember 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/023,502
; FILING DATE: August 6, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,315
; FILING DATE: No. 6015939ember 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: 119-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1589
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; US-08-747-574-2

Query Match 14.4%; Score 34.2; DB 3; Length 1589;
Best Local Similarity 49.2%; Pred. No. 1.5;
Matches 90; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 3 TTGTGATACGTTGGAAAGGGATATTGATTGGGAATGGGATGAAGATGAAGCGTAGATGA 62
DB 1205 TGTTCAGAGGTTGGAGAGAAAGTGAAGAGGAGGAGGAGGATCAATAAAGAGTTGA 1264
QY 63 GTTGGGAAAAAGTGTAGAACATATCGTGAAGATGAACCGTATTGTTGAAAAATCT 122
DB 1265 GGAGATAGAAGAGGATGAAGAGGTTGAGATAGAAGATCAACCTTATTTCAGTAACT 1324
QY 123 ACTGGTTAACTCAATCAGGAAAAACATCAAGAACTTAGGAAGTTGATTTCAGTTGAT 182
DB 1325 GTTGAAGGTTTAAAGACTCCACGAGATGAAGAGACITCTTAAGAGAGCTGAGCAA 1384
QY 183 AGA 185
DB 1385 AGA 1387

RESULT 12

US-08-913-842-12/c
; Sequence 12, Application US/08913842
; Patent No. 6028250
; GENERAL INFORMATION:
; APPLICANT: OHBA, Toshiharu
; APPLICANT: TAKAHASHI, Shuichi
; APPLICANT: ANMA, Yoshiko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GENE
; TITLE OF INVENTION: EXPRESSION USING SAID PROMOTER
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 7th Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,842
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 07-073043
; FILING DATE: 30-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/00777
; FILING DATE: 28-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: OHBA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-913-842-12

Query Match 14.3%; Score 34; DB 3; Length 1068;
Best Local Similarity 51.3%; Pred. No. 1.6;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 1 AATTGGATACGTTGGAAAGGGATATTGATTGGGAATGGGATGAAGATGAAGCGTAGAT 60
DB 1000 AATTCAATTCATAAATAAAGTGGCAATGGATAGATTTCACGTGATATAAAATTTGACAT 941
QY 61 GAGTTGGGAAAAAGTGTAGAACATATCGTGAAGATGAACCGTTATTGTTGAAAAAT 120
DB 940 GAAAGTGACAACATCTGTTGTAAGATGAAGTAAATGTAACAGAGATATGTGAAAAATTTAAATCT 881
QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAG 154
DB 880 CGCGGCTCTACTGCTTCATTCAGGGAAGTTGAGG 847

RESULT 13

US-08-257-073-2
; Sequence 2, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo

APPLICANT: de Taisne, Charles
APPLICANT: Time, John A.
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,305
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,183
FILING DATE: 20-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2981 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-257-073-2

Query Match 13.9%; Score 33; DB 1; Length 2981;
Best Local Similarity 60.7%; Pred. No. 3.8;
Matches 54; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 98 ATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATGAAGAAG 157
DB 42 ATTACAAAAATGTTATATAATGTACAGGAGAAAGTCAACAGGTAATACAGGAGGAGG 101

QY 158 TTAGGAAGTTGATTTATCAGTTGATAGAG 186
DB 102 TCAAGCAGGTAATACAGGAGGAGGTCAG 130

RESULT 14
US-08-184-009-119
Sequence 119, Application US/08184009
Patent No. 5833975
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,009
FILING DATE: 19-JAN-1994

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS

INFORMATION FOR SEQ ID NO: 119:
SEQUENCE CHARACTERISTICS:
LENGTH: 2981 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-184-009-119

Query Match 13.9%; Score 33; DB 2; Length 2981;
Best Local Similarity 60.7%; Pred. No. 3.8;
Matches 54; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 98 ATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGAGAAACATGAAGAAG 157
DB 42 ATTACAAAAATGTTATATAATGTACAGGAGAAAGTCAACAGGTAATACAGGAGGAGG 101

QY 158 TTAGGAAGTTGATTTATCAGTTGATAGAG 186
DB 102 TCAAGCAGGTAATACAGGAGGAGGTCAG 130

RESULT 15
US-08-458-356-119
Sequence 119, Application US/08458356
Patent No. 5942235
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,356
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2530

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

TELEX: 425066CURTMS

INFORMATION FOR SEQ ID NO: 119:

SEQUENCE CHARACTERISTICS:

LENGTH: 2981 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-458-356-119

Query Match 13.9%; Score 33; DB 2; Length 2981;

Best Local Similarity 60.7%; Pred. No. 3.8;

Matches 54; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 98 ATGAACCGTTATGTTGAAAAATCTACTGGTTAACTTCAATCAGGGAAAAACATGAAGAAG 157

Db 42 ATTTACAAAAATGTTTATAAATGTACAGGAGAAAGTCAACACAGGTAATACAGGAGGAGG 101

Qy 158 TTAGGAAGTTGATTATCATCAGTTGATAGAG 186

Db 102 TCAAGCAGGTAATACAGGAGGAGGTCAG 130

Search completed: November 15, 2003, 08:09:42

Job time : 43.0704 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 00:09:34 ; Search time 134.775 Seconds
(without alignments)
5771.357 Million cell updates/sec

Title: US-09-928-457-96

Perfect score: 238

Sequence: 1 AATTGGATACGTTGGAAAA.....ATGAAACTGGAACCAATT 238

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubnpa/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubnpa/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubnpa/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubnpa/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubnpa/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubnpa/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubnpa/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubnpa/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubnpa/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubnpa/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubnpa/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubnpa/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubnpa/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubnpa/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubnpa/US50_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubnpa/US50_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238	100.0	238	10	US-09-928-457-96
2	43	18.1	11670	12	US-10-240-452-26
3	39.2	16.5	6077	12	US-10-311-455-1364
4	38.4	16.1	3673778	12	US-10-312-841-1
5	37.4	15.7	4276	14	US-10-198-846-14002
6	37.4	15.7	11398	12	US-10-311-455-2411
7	37.4	15.7	3673778	12	US-10-312-841-2
8	36.8	15.5	664	12	US-09-814-353-17717
9	36.8	15.5	760	12	US-09-814-353-21458
10	36.8	15.5	760	14	US-10-198-846-13751
11	36.6	15.4	549	10	US-09-867-701-1884
12	36.2	15.2	5928	12	US-10-311-455-2212
13	36.2	15.2	11260	12	US-10-240-453-28
14	36.2	15.2	11260	14	US-10-239-676-20
15	36	15.1	14112	12	US-10-311-455-1416
16	35.8	15.0	617	10	US-09-917-800A-131

17	35.6	15.0	1440	12	US-09-882-227-65	Sequence 65, Appl
18	35.4	14.9	6029	12	US-10-311-455-1965	Sequence 1965, Ap
19	35.4	14.9	6954	12	US-10-311-455-1364	Sequence 1364, Ap
20	35.4	14.9	11398	12	US-10-311-455-2412	Sequence 2412, Ap
21	34.8	14.6	612	12	US-09-814-353-5041	Sequence 5041, Ap
22	34.8	14.6	612	12	US-09-814-353-11333	Sequence 11333, A
23	34.6	14.5	286	10	US-09-878-574-14604	Sequence 14604, A
24	34.6	14.5	3460	9	US-09-739-254-44	Sequence 44, Appl
25	34.6	14.5	3460	9	US-09-904-615-44	Sequence 44, Appl
26	34.6	14.5	3460	12	US-10-055-098-44	Sequence 44, Appl
27	34.6	14.5	3460	14	US-10-054-988-44	Sequence 44, Appl
28	34.6	14.5	6020	12	US-10-204-708-8	Sequence 8, Appl
29	34.6	14.5	6020	12	US-10-311-455-172	Sequence 172, App
30	34.6	14.5	11036	12	US-10-240-453-137	Sequence 137, App
31	34.6	14.5	11036	14	US-10-239-676-117	Sequence 117, App
32	34.4	14.5	5152	12	US-10-204-708-73	Sequence 73, Appl
33	34.4	14.5	6641	12	US-10-311-455-288	Sequence 288, App
34	34.4	14.5	6641	12	US-10-240-453-36	Sequence 36, Appl
35	34.4	14.5	397658	10	US-09-813-320-3	Sequence 3, Appl
36	34.2	14.4	1733	10	US-09-938-842A-3995	Sequence 3995, Ap
37	34.2	14.4	8013	12	US-10-311-455-1480	Sequence 1480, Ap
38	34.2	14.4	18598	12	US-10-311-455-360	Sequence 360, App
39	34	14.3	1135	12	US-10-027-632-119273	Sequence 119273,
40	34	14.3	1135	13	US-10-027-632-119273	Sequence 119273,
41	34	14.3	6917	12	US-10-311-455-2208	Sequence 2208, Ap
42	34	14.3	8530	12	US-10-311-455-405	Sequence 405, App
43	34	14.3	10133	12	US-10-311-455-432	Sequence 432, App
44	33.8	14.2	5396	12	US-10-240-483-153	Sequence 153, App
45	33.8	14.2	16724	12	US-10-311-455-1064	Sequence 1064, Ap

ALIGNMENTS

RESULT 1

US-09-928-457-96
; Sequence 96, Application US/09928457
; Patent No. US20020164603A1

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: DNA, specific proteins and peptides

TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
for obtaining them and their biological application.

NUMBER OF SEQUENCES: 99

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30 (OEB)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/928,457

FILING DATE: 2001-08-14

PRIOR APPLICATION DATA: 09/214,759

FILING DATE: 199-12-10

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-928-457-96

Query Match

Best Local Similarity 100.0%; Score 238; DB 10; Length 238;

Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATTGGATACGTTGGAAGGATATTTCATTGGGAATGGGATGAAGATAGCGTAGAT 60

1 AATTGGATACGTTGGAAGGATATTTCATTGGGAATGGGATGAAGATAGCGTAGAT 60

```
QY 61 GAGTTGGGAAAAAGTGTAGACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 120
Db 61 GAGTTGGGAAAAAGTGTAGACATATCGTAAAGATGAACCGTTATTGTTGAAAAAT 120
QY 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 180
Db 121 CTACTGGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTG 180
QY 181 ATAGAGTTAGATTTCTCGAACTTTTGTGAGGATTTCTATCAAAAACCTGGAACCAATT 238
Db 181 ATAGAGTTAGATTTCTCGAACTTTTGTGAGGATTTCTATCAAAAACCTGGAACCAATT 238

RESULT 2
US-10-240-452-26
; Sequence 26, Application US/10240452
; Publication No. US20030162194A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Apoptosis
; FILE REFERENCE: 5013.1006
; CURRENT APPLICATION NUMBER: US/10/240.452
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03969
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 78
; SEQ ID NO 26
; LENGTH: 11670
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-452-26

Query Match 18.1%; Score 43; DB 12; Length 11670;
Best Local Similarity 49.3%; Pred. No. 0.8;
Matches 112; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 3 TTGGATACGTTGGAAGGGATATTTCATTGGGAATCGGATGAAGATAACCGTAGATGA 62
Db 265 TTAGGTTATGATGGAATAATATTTTATATGGAATTATATTAATAAATTTTGAAGTGGG 324
QY 63 GTTGGGAAAAAGTGTAGACATATCGTAAAGATGAACCGTTATTGTTGAAAAATCT 122
Db 325 GTAAGAAATATAAATGTTTTTATTATAGATTATTAATAAATTTTGTGTTTGGATA 384
QY 123 ACTGTTAACTTCAATCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTGAT 182
Db 385 GAGGTTGATTTTAGATCGTAAATAATAGTAATTAATAATTTTAAATTTATGATTAGT 444
QY 183 AGAGTTAGATTTCTCGAACTTTTGTGAGGATTTCTATCAAAAACCTG 229
Db 445 AGAGTTAGGCTGTGTTGTTTATATTTTATATTTTATTAATGTTG 491

RESULT 3
US-10-311-455-2191
; Sequence 2191, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
```

```
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2191
; LENGTH: 6077
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2191

Query Match 16.5%; Score 39.2; DB 12; Length 6077;
Best Local Similarity 50.5%; Pred. No. 5.8;
Matches 95; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 24 ATATTGATTGGGAATGGATGAAGATGAACGCTAGATGAGTTGGGAAAAAGTGTAGA 83
Db 824 ATAGTGGTTTGGGGTGGAAAGTTTTTAAATGGAATGTTGTTGGTGTGATGGA 883
QY 84 ACATATCGGTAAGAATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGG 143
Db 884 GTATTTAATTTTGGTTTAGGAGTTTTTGAAGTTTTTTTGGTTAAGATGATTTCGAG 943
QY 144 AAAACATGAAGAAGTTAGGAAGTTGATTAATCAGTTAGTAGATTCTTCGAACT 203
Db 944 GAAATGATTGATTATTGTTGTTGTTTTTAAATGATGATATAGATTAGATTTTATAAT 1003
QY 204 TTTGTGAG 211
Db 1004 TATATGAG 1011

RESULT 4
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312.841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 16.1%; Score 38.4; DB 12; Length 3673778;
Best Local Similarity 49.5%; Pred. No. 93;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 25 TATTGATTGGGAATCGGATGAAGATGAACGCTAGATGAGTTGGGAAAAAGTGTAGA 84
Db 433638 TTTTATTATTTTGTGTTGTTGTTTTTGGGAACGATTTAATTCGTC 433697
QY 85 CATATCGGTAAGAATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGGA 144
```



```

; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17717
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-353-17717

Query Match 15.5%; Score 36.8; DB 12; Length 664;
Best Local Similarity 54.4%; Pred. No. 10;
Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 95 AGAATGAACCGTTATTGTTGAAAATCTACTGTTAACTTCAATCAGGGAACATGAAG 154
Db 627 ACAATCCAGTTTGTGAGAAAGAAACAATTCATCTATATCTGCAAAATGTAAT 568

QY 155 AAGTTAGGAAGTTGATTTATCAGTTGATAGATTCTGGAACCTTTGTGAGGGA 214
Db 567 ATTTAGGAATAGTCTCTAGTTAATAATGTTTACAGTTCTAATCATCTGTTTGA 508

QY 215 TTCTATGAAAACCTGG 230
Db 507 TTTGTGAGAAATGG 492

RESULT 9
US-09-814-353-21458/c
; Sequence 21458, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; APPLICANT: Lee, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21

```

```

; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21458
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168,
; LOCATION: 169, 170, 171, 172, 173, 174, 175
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353-21458

Query Match 15.5%; Score 36.8; DB 12; Length 760;
Best Local Similarity 54.4%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 95 AGAATGAACCGTTATTGTTGAAAATCTACTGTTAACTTCAATCAGGGAACATGAAG 154
Db 522 ACAATCCAGTTTGTGAGAAAGAAACAATTCATCTATATCTGCAAAATGTAAT 463

QY 155 AAGTTAGGAAGTTGATTTATCAGTTGATAGATTCTGGAACCTTTGTGAGGGA 214
Db 462 ATTTAGGAATAGTCTCTAGTTAATAATGTTTACAGTTCTAATCATCTGTTTGA 403

QY 215 TTCTATGAAAACCTGG 230
Db 402 TTTGTGAGAAATGG 387

RESULT 10
US-10-198-846-13751/c
; Sequence 13751, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13751
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-198-846-13751

Query Match 15.5%; Score 36.8; DB 14; Length 760;
Best Local Similarity 54.4%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 95 AGAATGAACCGTTATTGTTGAAAATCTACTGTTAACTTCAATCAGGGAACATGAAG 154
Db 522 ACAATCCAGTTTGTGAGAAAGAAACAATTCATCTATATCTGCAAAATGTAAT 463

QY 155 AAGTTAGGAAGTTGATTTATCAGTTGATAGATTCTGGAACCTTTGTGAGGGA 214
Db 462 ATTTAGGAATAGTCTCTAGTTAATAATGTTTACAGTTCTAATCATCTGTTTGA 403

QY 215 TTCTATGAAAACCTGG 230
Db 402 TTTGTGAGAAATGG 387

RESULT 11

```


US-09-867-701-1884/c

; Sequence 1884, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1884
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(549)
; OTHER INFORMATION: n = A,T,C or G

US-09-867-701-1884

Query Match 15.4%; Score 36.6; DB 10; Length 549;
Best Local Similarity 55.2%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 106 TTATTTGTAATAATCTACTGGTTAACTCAATCAGGAGAAACATGAAGAGTTAGGAAG 165
Db 335 TTGTGAGAAGAACTAATCTACTCTATTTCTGCAATGTAATNNTAGGAA 276
QY 166 TTGATTATCAGTTGATAGATAGATTTTCTGGAACCTTTTGTGAGGATCTATGAAAA 225
Db 275 TTAGTCTAGGTTAATAATGTTTACAGTCTTAACATCTGTCTTTAGATTTCGTGAAA 216
QY 226 ACTGG 230
Db 215 ATGG 211

RESULT 12

US-10-311-455-2212
; Sequence 2212, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2212
; LENGTH: 5928
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-2212

Query Match 15.2%; Score 36.2; DB 12; Length 5928;
Best Local Similarity 53.1%; Pred. No. 33;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 25 TATTTGATTTGGGAATCGGATGAAGATAACCGTAGATGAGTTGGGAAAAAAGTGTAGAA 84
Db 5179 TAATTTAATATATATGTAATTTGATGCTATATTAGTAGAGAGAGAGAAAAATTTTTTAAA 5238
QY 85 CATATCGGTAAAGAATGAACCGTTATTGTTGAAAAATCTACTGTTAACTTCAATCAGGA 144
Db 5239 GAATATATTTCATTAAGAATGAGTATTGAGTAATAATTATTAGATTTCATGAGTTGGT 5298
QY 145 AAACATGAAGAAGTTAGGAAGTTGA 169
Db 5299 GAAAGCTATAAAATTTATAAAATTTA 5323

RESULT 13

US-10-240-453-28
; Sequence 28, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: By Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 28
; LENGTH: 11260
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-453-28

Query Match 15.2%; Score 36.2; DB 12; Length 11260;
Best Local Similarity 48.3%; Pred. No. 42;
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 18 AAAGGGATATTTGATTGGGAATGGGATGAAGATAGCGTAGATGAGTTGGGAAAAAAGT 77
Db 9770 AAAGTGTATATTAATGCTGTGTAAGAGTAGAATAATAATAATTTTTTTTTCAGAAAAA 9829
QY 78 GTTAGAACATATCGGTAAAGATGAACCGTTATTGTTGAAAAATCTACTGTTAACTCAA 137
Db 9830 ATAATAAAAAATATATTAGGAAATTAGGAATAAATTTAAGTAATAATAATAATTTTAA 9889
QY 138 TCAGGGAAACATGAAGAGTTAGGAAGTTAGTATTTATCAGTTGATAGAGTTTCT 197
Db 9890 TAGAGTAATTAATGAAGTAGTTATTGTTAGATATATATAAGAAAAAGTGTATTTT 9949
QY 198 GGAACATTTTGTGAGGGAATCTATGAAAAA 226
Db 9950 GGAATCTTTTAAATGTTTGTGGGTTAA 9978

RESULT 14

US-10-239-676-20
; Sequence 20, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:

APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 20
LENGTH: 11260
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

Query Match 15.2%; Score 36.2; DB 14; Length 11260;
Best Local Similarity 48.3%; Pred. No. 42;
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 18 AAGGATATTTGATTTGGGAAATGGGATGAAGATAGCGTAGATGAGTTGGGAAAAAAGT 77
DB 9770 AAAGTGTATATTAATGCTGTGTAAGAAAGTAGAATAATAATTTTTTTTTCAGAAAAA 9829
QY 78 GTTAGAACATATCGGTGAAGATGAACCGTTATTGTGAAAAATCTACTGTTAACTCAA 137
DB 9830 ATATTAATATATATAGGAATTTAGGAATTAATTAATTAATTAATTTTAAATTTA 9889
QY 138 TCAGGAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTGATAGATTAGATTTCT 197
DB 9890 TAGAGGTAATTTATGAAGTAGTTATTGCTAGATATATATAAGAAAAAGTGTTTTATTT 9949
QY 198 GGAACCTTTTGTAGGCGATCTATGAAAA 226
DB 9950 GGAATGTTATATGTTTGTGGGTTAA 9978

RESULT 15
US-10-311-455-1416
Sequence 1416, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Peter
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1416
LENGTH: 14112
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1416
Query Match 15.1%; Score 36; DB 12; Length 14112;
Best Local Similarity 49.0%; Pred. No. 51;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 13 TTGGAAAAAGGATATTTGATTTGGGAAATGGGATGAAGATAGCGTAGATGAGTTGGGAAAA 72
DB 1492 TTCAATATGGAGTTTATAATTTTCGTTTAAAGCTTTTAAAGGTTTGTGGGTAATTTGAGGAAA 1551
QY 73 AAAGTGTTAGAACATATCGGTGAAGATGAACCGTTATTGTGAAAAATCTACTGGTTAAC 132
DB 1552 AAAGGTGGTAGGTTTGGTTGAGAAATAGGGTTTTTTTTTAATATTCGGTAGTAGG 1611
QY 133 TTCAATTCAGGAAAAACATGAAGAAGTTAGGAAGTTGATTTATCAGTTGATAGATTAGAT 192
DB 1612 GGTTCGTAGCTGTAAGATAGAGAGGATGGGTTTCGTTGTTTTTATTTTCGTTTAT 1671
QY 193 TTTCTGGAACCTTTTGT 208
DB 1672 TTTTTCGGGTTTGT 1687

Search completed: November 15, 2003, 08:32:46
Job time : 145.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:56:24 ; Search time 960.589 Seconds
(without alignments)
9241.609 Million cell updates/sec

Title: US-09-928-457-99

Perfect score: 217

Sequence: 1 AATTCGGACAGTATGATAC.....ATGACGATGACAGATTAAATT 217

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: gb.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.rod.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	100.0	217	1	AF169477
2	217	100.0	217	6	A68928
3	217	100.0	217	6	BD063028
4	191.4	88.2	349061	1	NMA222491
5	130.4	60.1	858	6	AX024058
6	130.4	60.1	858	6	AX024111
7	42.8	19.7	223954	2	AC106472
8	42.8	19.7	261007	2	AC133419
9	41.4	19.1	163284	10	AC129204
10	41.4	19.1	183202	10	AC125488
11	40.4	18.6	181815	2	AC131594
12	40.4	18.4	301450	1	AF003188
13	39.2	18.1	695	3	AF412737
14	39.2	18.1	695	3	AF412737
15	39.2	18.1	163453	2	EX088572
16	38.8	17.9	153336	2	AC025020
17	38.8	17.3	156277	9	AC107015
18	38.8	17.9	163022	2	AC016851
19	38.8	17.9	163737	9	AC108697
20	37.6	17.3	602	6	AX381792
21	37.6	17.3	3507	9	AF077367
22	37.6	17.3	3574	9	BC018122
23	37.6	17.3	3615	9	D88687
24	37.6	17.3	3817	6	BD094040
25	37.6	17.3	3826	9	AR274914
26	37.6	17.3	3826	9	HSTHIORED
27	37.6	17.3	3826	9	S79851
28	37.6	17.3	83755	9	AC090107
29	37.6	17.3	114103	5	AL672171
30	37.6	17.3	251845	2	AC097949
31	37.6	17.3	303250	1	AF003193
32	37.4	17.2	173929	9	AC007217
33	37.4	17.1	38432	8	SPAC11E3
34	36.8	17.0	235700	2	AC103577
35	36.6	16.9	177987	9	AC107052
36	36.2	16.7	311	3	HHU32460
37	36.2	16.7	155568	2	AP053393
38	36.2	16.7	238256	2	AC094282
39	36.2	16.7	242935	2	AC094306
40	36	16.6	2000	6	AX509017
41	36	16.6	2003	6	AX461377
42	36	16.6	2004	6	AX461324
43	36	16.6	3390	9	HSTHIORED
44	36	16.6	102873	8	ATF26GS
45	36	16.6	110000	6	BD061520_5

ALIGNMENTS

RESULT 1
AF169477
LOCUS AF169477 217 bp DNA linear BCT 09-AUG-2000
DEFINITION Neisseria meningitidis strain 22491 clone Em10 unknown sequence.
ACCESSION AF169477
VERSION AF169477.1 GI:9754685
KEYWORDS Neisseria meningitidis
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 217)
Perrin, A., Nassif, X., and Tinsley, C.R.
AUTHORS Identification of regions of the chromosome of Neisseria
TITLE meningitidis and Neisseria gonorrhoeae which are specific to

Pred. No. is the number of results predicted by chance to have a

```

pathogenic Neisseriae
Unpublished
2 (bases 1 to 217)
Perrin,A., Nassif,X. and Tinsley,C.R.
Direct Submission
TITLE
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugivard, Paris 75015, France
LOCATION/Qualifiers
1. .217
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:487"
/clone="Em10"
76 a 31 c 44 g 66 t

BASE COUNT
ORIGIN
Query Match 100.0%; Score 217; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60
DB 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60

QY 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120
DB 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120

QY 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180
DB 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180

QY 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217
DB 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217

RESULT 2
LOCUS A68928 217 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 99 from Patent WO9802547.
ACCESSION A68928
VERSION A68928.1 GI:4759847
KEYWORDS
SOURCE unidentifed
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 217)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and Merker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
JOURNAL
COMMENT
PATENT: WO 9802547-A 99 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
Other publication FR 2751000 19980116.
FEATURES
source
1. .217
/organism="unidentifed"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
76 a 31 c 44 g 66 t

BASE COUNT
ORIGIN
Query Match 100.0%; Score 217; DB 6; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60
DB 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60

pathogenic Neisseriae
Unpublished
2 (bases 1 to 217)
Perrin,A., Nassif,X. and Tinsley,C.R.
Direct Submission
TITLE
JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
rue de Vaugivard, Paris 75015, France
LOCATION/Qualifiers
1. .217
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:487"
/clone="Em10"
76 a 31 c 44 g 66 t

BASE COUNT
ORIGIN
Query Match 100.0%; Score 217; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60
DB 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60

QY 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120
DB 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120

QY 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180
DB 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180

QY 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217
DB 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217

RESULT 2
LOCUS A68928 217 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 99 from Patent WO9802547.
ACCESSION A68928
VERSION A68928.1 GI:4759847
KEYWORDS
SOURCE unidentifed
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 217)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and Merker,P.
TITLE DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
JOURNAL
COMMENT
PATENT: WO 9802547-A 99 22-JAN-1998;
INST NAT SANTE RECH MED (FR)
Other publication FR 2751000 19980116.
FEATURES
source
1. .217
/organism="unidentifed"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
76 a 31 c 44 g 66 t

BASE COUNT
ORIGIN
Query Match 100.0%; Score 217; DB 6; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60
DB 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60

```

```

QY 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120
DB 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120

QY 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180
DB 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180

QY 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217
DB 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217

RESULT 3
LOCUS BD063028 217 bp DNA linear PAT 27-AUG-2002
DEFINITION DNA and specific proteins or peptides of the Neisseria meningitidis
applications.
ACCESSION BD063028
VERSION BD063028.1 GI:22608631
KEYWORDS JP 2001504684-A/90.
SOURCE unidentifed
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 217)
AUTHORS Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and Merker,P.
TITLE DNA and specific proteins or peptides of the Neisseria meningitidis
applications
JOURNAL Patent: JP 2001504684-A 90 10-APR-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX
PLANK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,
SMITHKLINE BEECHAM
COMMENT PN JP 2001504684-A/90
PD 10-APR-2001
PF 11-JUL-1997 JP 1998505685
PR 12-JUL-1996 FR 96/08768
PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI
CARLA VINALS,
PI PETRA MERKER
PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/68, G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1. .217
/organism="unidentifed"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
76 a 31 c 44 g 66 t

BASE COUNT
ORIGIN
Query Match 100.0%; Score 217; DB 6; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60
DB 1 AATTCGGACAGTATGAATACACGGGATTAATACAGGTAAGTTTCATTACACGGAAGAAC 60

QY 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120
DB 61 CTTTAAAGAAATATGAAGGTATTACCTGTTTGCACACGGGAATGCTAAATATGCC 120

QY 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180
DB 121 GAGTTTTTCACTGAATACAGCAATCAGCCATTTCATTATATTTGACTGGATGCTGAA 180

QY 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217
DB 181 TGTGGACTTTATAGATAATGACGATGAAGATTAAAT 217

```

```

RESULT 4
NMA222491
LOCUS
DEFINITION
  NMA222491 Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 2/7.
ACCESSION
  AL162753 AL157959
VERSION
  AL162753.2 GI:7379120
KEYWORDS
  .
SOURCE
  Neisseria meningitidis Z2491
  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
  Neisseriaceae; Neisseria.
  1 (bases 1 to 349061)
  Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C.,
  Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T.,
  Davies, R.M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N.,
  Holroyd, S., Jagers, K., Leather, S., Moule, S., Mungall, K.,
  Quail, M.A., Rajandream, M.A., Rutherford, K.M., Simmonds, M.,
  Skelton, J., Whitehead, S., Spratt, B.G. and Barrell, B.G.
  Complete DNA sequence of a serogroup A strain of Neisseria
  meningitidis Z2491
  Nature 404 (6777), 502-506 (2000)
JOURNAL
  Nature 404 (6777), 502-506 (2000)
MEDLINE
  20222555
PUBMED
  10761915
REFERENCE
  2 (bases 1 to 349061)
  Parkhill, J.
  Direct Submission
  Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
  sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
  Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
  Notes:
  Details of N. meningitidis sequencing at the Sanger Centre are
  available on the World Wide Web.
  (URL, http://www.sanger.ac.uk/projects/N_meningitidis/).
  Location/Qualifiers
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
FEATURES
  source
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
  repeat_unit
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
  misc_feature
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
  misc_feature
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
  misc_feature
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmembrane domains"
      /codon_start=1
      /product="putative integral membrane protein"
      /protein_id="CAB83669.1"
      /db_xref="GI:7379121"
      /db_xref="SPTREMBL:Q9JWH7"
      /translation="MQQNRKSSPPIVNLIVSVLWIASLVNVAFLVGHGSMGLTY
      LILGSIPLSDIRYCAVYANYVLAIAVLLALRKVPVYHAAPFWGLAVAFVSKAVVY
      DEAGNTSDIVRYAGFYLMWAAFAVASIGTFAGNKKERKAASAADGTKDV"
      671..1492
      /gene="hemK"
      671..1492
      /gene="hemK"
      /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
      to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
      scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
      PS00092 N-6 Adenine-specific DNA methylases signature"
      /codon_start=1
      /transl_table=11
  misc_feature
    1..349061
      /organism="Neisseria meningitidis Z2491"
      /mol_type="Genomic DNA"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      complement(24..206)
      /note="ATR repeat, hmms hit to HMV ATR (1 - 183), score:
      310.39"
      /label=ATR
      209..212
      220..681
      /gene="NMA0368"
      220..681
      /gene="NMA0368"
      /note="NMA0368, probable integral membrane protein, len:
      153 aa; contains four probable transmem
```


Bouchardon, A. and Renauld, M.G.
Patent: FR 2785293-A 54 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)
Location/Qualifiers
source 1. -858
/organism="Neisseria gonorrhoeae"
/mol_type="genomic DNA"
/db_xref="taxon:485"
CDS 1. -858
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC08931.1"
/db_xref="GI:10184423"
/translation="MSEELKMSFPTVIEHLGVKMYGHTVPAIAELIANAYDACATE
VEVRLFDKEHIVIKDNGIGMSFDEINDFYLRGNREPKQASPCGRIPITGKGLG
KLALPRLGNKISITOGNERVTFILDAIEIKSRIQPERKQSPKNTENGTTIT
LHELKQGYPIDINVGHSRLFDPAQDFKIKVLSNGSEPRIDNGLKYNLVTPOFE
WEYQDLATNLSLSKFEQREYSGLIQGFITTEKPELKNMKGITLFGANGRMVNPEF
FTDSSESHF"
BASE COUNT 302 a 151 c 178 g 227 t
ORIGIN
Query Match 60.1%; Score 130.4; DB 6; Length 858;
Best Local Similarity 97.5%; Pred. No. 2.4e-22;
Matches 154; Conservative 0; Mismatches 1; Indels 3; Gaps 2;
QY 1 AATCGACAGTATGAATACAGCGATTAAATCAAGGTAAGTTCATTACACGGAAAAAC 60
|||||
Db 701 AATTCGAACAGTATGAATACAGCGATTAAATCAAGGTAAGTTCATTACACGGAAAAAC 760
|||||
QY 61 CTTTAAAGTAATATGAAGTATACCTTTGTCGAACGG--GAATGTAATATGTC 118
|||||
Db 761 CTTTAAAGTAATATGAAGTATACCTTTGTCGAACGGAGCAATGCTAAATATGTC 820
|||||
QY 119 CCGAGTTTTCCTACTGAATAGCAATCCAGCCATTCTTA 156
|||||
Db 821 CCGAGTTTTCCTG-ATAGCGAATCCAGCCATTCTTA 857
|||||
RESULT 7
AC106472/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-15201, ***SEQUENCING IN PROGRESS
*** 3 unordered pieces.
ACCESSION AC106472
VERSION AC106472.4 GI:30580924
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 223954)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Banderet, A., Barker, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbini, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaekeleneh, O., Okwunnu, G., Olarunpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Popovic, D., Primus, E., Pu, L.,
Plummer, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, C., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 223954)
Worley, K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 223954)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23265983.
The sequence in this assembly is a combination of BAC-based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKX
Center clone name: CH230-15201
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 20381 bases at least Q40
Consensus quality: 20755 bases at least Q30
Consensus quality: 21032 bases at least Q20
Estimated insert size: 215302; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 221228: contig of 221228 bp in length
2 221328: gap of unknown length
3 221329: contig of 1096 bp in length
4 222426: gap of unknown length
5 222524: gap of unknown length
6 222525: contig of 1430 bp in length.

Location/Qualifiers
1. 223954
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-15201"
/note="clone boundary"
clone_end:596
site:EcoRI
end_sequence:BH335992"
1514..1860
/note="clone boundary"
clone_end:17
site:EcoRI
end_sequence:BH335441"
95383..97199
/note="wgs_end_extension"
clone_end:17"

BASE COUNT 62554 a 48124 c 45176 g 56913 t 11187 others
ORIGIN

Query Match 19.7%; Score 42.8; DB 2; Length 223954;
Best Local Similarity 53.6%; Pred. No. 0.63;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 42 TTCATTACAGGAAACCTTTAAGATATATGAAGGTATACCTTTTGCCCAAC 101
DB 131056 TTCATTGAGAGCATGCTATATAAACAAGATCAATTTATCATATTTCCAT 130997

QY 102 GGAATGTAAATATGCCGAGTTTTCACCTGAATAGCAATCCAGCATTTCTATTCAT 161
DB 130996 GAAATAGACAACTGCCCTTTTAAATAATAAATAGTCCATTTATTTTGT 130937

QY 162 ATTGACTGGATGGCTGAATGTGACCTTATAGATATGACATGA 207
DB 130936 TTCATTTTAAATTTAGTTTGGAGATCCATAGCTAATGAGTATGA 130891

RESULT 8
AC133419/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-52418, *** SEQUENCING IN PROGRESS
***, 4 unordered pieces.
ACCESSION
AC133419
VERSION
AC133419.2 GI:25007534
KEYWORDS
HTG, HTGS, PHASE1, HTGS, DRAFT, HTGS_ENRICHED.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 261007)
Muzny, D., Marie, Metaker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devilla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fraser, C., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frederick, S., Gant, A., Garcia, A., Garner, T., Garza, M.,
Gebregorjis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hernandez, J.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hines, M.,
Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hughes, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M., McNeill, T., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Pua, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smajda, D.,
Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D., Holt, R., Smith, H.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 261007)
Rat Genome Sequencing Consortium.
Submitted (12-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 261007)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:22795100.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine


```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAYS
Center clone name: CH230-52418
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 223700 bases at least Q40
Consensus quality: 226843 bases at least Q30
Consensus quality: 228976 bases at least Q20
Estimated insert size: 228564; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 240201: contig of 240201 bp in length
240202 240301: gap of unknown length
240302 256939: contig of 16638 bp in length
256940 257039: gap of unknown length
257040 258335: contig of 1496 bp in length
258336 258635: gap of unknown length
258636 261007: contig of 2372 bp in length.
-----
FEATURES
         source
         1..261007
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-52418"
         1..1542
            /note="wgs contig"
         1593..4461
            /note="wgs contig"
         252331..253555
            /note="wgs contig"
         255445..256939
            /note="wgs contig"
BASE COUNT 68651 a 52580 c 49193 g 59942 t 30641 others
ORIGIN
Query Match      19.7%; Score 42.8; DB 2; Length 261007;
Best Local Similarity 53.6%; Pred. No. 0.61;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 42 TTCAATCAACGGAACCTTTAAAGAAATATATGAAGGTATACCTTGTGGCCAC 101
Db 95269 TTCAATTTGGAGAGAGATGCTATATAACAAAGATCAAAATTTACTATCATATTTCCACAT 95210

QY 102 GGAATGTAATAATGCCCAGTGTTCCTCACTGAATAGCAATCCAGCAATTTCTATTCAT 161
Db 95209 GAAATAGACAAACTGCCCACTTTTAAATAATAAATAGTCCATTTATTTTATTTGT 95150

QY 162 ATTGACTGGATGCTGAATGTGACTTTATAGATATAGCATGA 207
Db 95149 TTTCAATTTTAAATTTAGTTTGGAGAAATCCCATAGTAATAGTATGA 95104

RESULT 9
AC129204
LOCUS AC129204 163284 bp DNA linear ROD 23-FEB-2003
DEFINITION Mus musculus chromosome 12 clone RP24-230P5, complete sequence.
ACCESSION AC129204
VERSION AC129204.4 GI:28475615
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 163284)
McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163284)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 163284)
McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 163284)
McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Feb 23, 2003 this sequence version replaced gi:28372744.
-----
FEATURES
         Location/Qualifiers
         source
         1..163284
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="12"
            /clone="RP24-230P5"
BASE COUNT 52814 a 29602 c 29566 g 51302 t
ORIGIN
Query Match      19.1%; Score 41.4; DB 10; Length 163284;
Best Local Similarity 56.1%; Pred. No. 1.5;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 27 TTAATCAAGGTAAAGTTCATTACACGGAACCTTTAAAGAAATATATGAAGGTATT 86
Db 148567 TTAATAATAGTCACATTCATTAAAGCTTAGTACCCTTGATATTAACCAATATCTT 148626

QY 87 ACCTTCTTTGCCAACGGGAATGGTAAATATGCCGAGTTTTTCACTGAATAGCATCA 146
Db 148627 TCCATATTTTGTAGTTTACTCAATATATATAGGTTTCAATGAAACAAGTATCA 148696

QY 147 GCATTTCTATTCAATTTT 165
Db 148687 TACAATTAATAACAATT 148705

RESULT 10
AC125488
LOCUS AC125488 183202 bp DNA linear ROD 21-DEC-2002
DEFINITION Mus musculus chromosome 12 clone RP23-75M24, complete sequence.
ACCESSION AC125488
VERSION AC125488.3 GI:27356773
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 183202)
McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 183202)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (27-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 183202)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (06-DEC-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 183202)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (21-DEC-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Dec 21, 2002 this sequence version replaced gi:26080599.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0075M24

----- Location/Qualifiers -----
1. 183202
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/clone="RP23-75M24"
60541 a 33423 c 33788 g 55450 t

BASE COUNT
ORIGIN

Query Match 19.1%; Score 41.4; DB 10; Length 183202;
Best Local Similarity 56.1%; Pred. No. 1.4;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 27 TTATACAGGTAAGTTCATTACACGGAAAACCTTTAAAGATATATGAAGGTATT 86
Db 12250 TTAATAATAGTGACATTCATTAAAGCTTAGTACCCCTTGGATATTAAACACATATCTT 12309
QY 87 ACCTTGTTTGGCAACGGAAATGGTAAATATGCCGAGTTTTCACCTGAATAGCAATCCA 146
Db 12310 TCCATATTGTTAGTTATACATATATATAGTTTACATGAACAGTATTCA 12369
QY 147 GCCATTTCATCATATT 165
Db 12370 TACAATTATTATACATTT 12388

RESULT 11
AC129541/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC129541
Mus musculus clone RP23-248M12, *** SEQUENCING IN PROGRESS ***, 4
unordered pieces.
AC129541
GI:30984790
HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198303)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-248M12
Unpublished
2 (bases 1 to 198303)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,

Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kells,C., Lander,E., Levine,T., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,I., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,S., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 198303)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cooke,P., Cooke,K., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hagos,B., Haploplan,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2003 this sequence version replaced gi:29126476.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Smit, A.F.A. & Green, P. (1996-1997)
----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L26920
Center clone name: 248_M_12

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved
* 1 82226: contig of 82226 bp in length
* 82227: gap of 100 bp
* 82227 112802: contig of 30476 bp in length
* 112803 112902: gap of 100 bp
* 112903 188123: contig of 75221 bp in length
* 188124 188223: gap of 100 bp
* 188224 198303: contig of 10080 bp in length.
Location/Qualifiers
1. 198303
/organism="Mus musculus"

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-248M12"
/clone_lib="RP23-248M12"
/collection="Female Mouse BAC"
BASE COUNT 63714 a 36811 c 36425 g 60546 t 807 others
ORIGIN
Query Match 18.9%; Score 41; DB 2; Length 198303;
Best Local Similarity 49.5%; Pred. No. 1.8;
Matches 104; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 6 GCACAGTATGATACAGCGGATTAATACAGGTAAAGTTCATTACACGCGAAACCTTTA 65
Db 155858 GCGAAGATATATAACCTCAGCTGCTGAATATAAAAGCAGGAAAAACAAACAAACA 155799
QY 66 AGAATATATGAAGAGTATACCTTTGTCACCGGAATGGAATATATGCCGAGTT 125
Db 155798 AAACAAAAAANAANCAAAATTAATACAGAAATAAATGAAACAAATAGCTGGTT 155739
QY 126 TTTCAGTGAATAGCGAATCCAGCCATTTCTATTCATATTCGCTGGATGCTGATGG 185
Db 155738 TTTCAGAGATGAATAGATTGAAATTTTATCCAAATTTACAAATTTGCGAGAAAGAGA 155679
QY 186 ACTTTATGATATGACGATGAGATTTAA 215
Db 155678 AGATCCAAATTAATGAATTCAGATGAAA 155649

RESULT 12
AC131594 181815 bp DNA linear HTG 26-FEB-2003
LOCUS Mus musculus clone RP24-568P18, WORKING DRAFT SEQUENCE, 4 unordered
DEFINITION pieces.
ACCESSION AC131594
VERSION AC131594.5 GI:28565729
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Biren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-568P18
Unpublished
2 (bases 1 to 181815)
Biren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B.,
Camata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Minova,T., Mienga,V.,
Murphy,I., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 181815)
Biren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Archchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhalter,B., Camata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,

```

```

Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Minova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 26, 2003 this sequence version replaced gi:28412028.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seg.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26731
Center clone name: 568_P_18
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181080 bases at least Q40
Consensus quality: 181410 bases at least Q30
Consensus quality: 181465 bases at least Q20
Insert size: 184000; agarose-fp
Insert size: 181515; sum-of-contrigs
Quality coverage: 9.0 in Q20 bases; agarose-fp
Quality coverage: 9.2 in Q20 bases; sum-of-contrigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 35214: contig of 35214 bp in length
* 35215 35314: gap of 100 bp
* 35315 60599: contig of 25285 bp in length
* 60600 60699: gap of 100 bp
* 60700 97829: contig of 37130 bp in length
* 97830 97929: gap of 100 bp
* 97930 181815: contig of 83886 bp in length.
Location/Qualifiers
1..181815
/oranism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-568P18"
/clone_lib="RP24-568P18"
/misc_feature 1..35214
/notes="assembly_fragment
clone_end:56
vector_side:left"
misc_feature 35315..60599
/notes="assembly_fragment"
misc_feature 60700..97829
/notes="assembly_fragment"
misc_feature 97930..181815
/notes="assembly_fragment"
FEATURES
source
1..181815
/oranism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-568P18"
/clone_lib="RP24-568P18"
/misc_feature 1..35214
/notes="assembly_fragment
clone_end:56
vector_side:left"
misc_feature 35315..60599
/notes="assembly_fragment"
misc_feature 60700..97829
/notes="assembly_fragment"
misc_feature 97930..181815
/notes="assembly_fragment"

```

```

clone_end:T7
vector_side:right
BASE COUNT 58499 a 33923 c 33630 g 55462 t 301 others
ORIGIN

Query Match 18.6%; Score 40.4; DB 2; Length 181915;
Best Local Similarity 49.5%; Pred. No. 2.6;
Matches 104; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 6 GGACAGTATGAATCAGCGGATTAAATACAGCTAAGTTCATTACAAACGGAACACCTTTA 65
DB 173210 GGAAGAAATATAACCTCGGTCTGAAATATAAAGACGGAACAAACAAACAA 173269

QY 66 AAGATAATATGAAGGTATTACCTTGTTCACACGGGAATGTAATATGCCGAGTT 125
DB 173270 AAACAAAAAACCACCAAAATAATTACAAAGATAAATGAACAAATAGCTGGTT 173329

QY 126 TTTCACGAATAGCAATCCAGCCATTCTATTCTATTTGACATGCGTGAATCTGG 185
DB 173330 TTTCAGAAATGAATAGATTGAAAATCTTATCCAAATTACAAATGGCAGAAAGA 173389

QY 186 ACTTTATAGATAATCAGCATGAAGATTAA 215
DB 173390 AGATCCAAATTAATGAATTCAGATGAAA 173419

RESULT 13
AP003188 301450 bp DNA linear BCT 10-JUL-2002
LOCUS Clostridium perfringens str. 13 DNA, complete genome, section 4/10.
DEFINITION AP003188 BA000016
ACCESSION AP003188.2 GI:18146726
VERSION
KEYWORDS
SOURCE Clostridium perfringens str. 13
ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1
AUTHORS Shimizu,T., Ohtani,K., Hirakawa,H., Ohshima,K., Yamaeshita,A.,
Shiba,T., Ogasawara,N., Hattori,M., Kuhara,S. and Hayashi,H.
TITLE Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (2), 996-1001 (2002)
MEDLINE 21664373
PubMed 11792842
REFERENCE 2 (bases 1 to 301450)
AUTHORS Shimizu,T.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2001) Tohru Shimizu, Institute of Basic Medical
Sciences, University of Tsukuba, Department of Microbiology; 1-1-1
Tenohdai, Tsukuba, Ibaraki 305-8575, Japan
(E-mail: tshimizu@md.tsukuba.ac.jp, Tel:81-298-53-3354,
Fax:81-298-53-3354)
COMMENT On Jan 14, 2002 this sequence version replaced gi:18144387.
FEATURES
Location/Qualifiers
1..301450
/organism="Clostridium perfringens str. 13"
/mol_type="Genomic DNA"
/strain="13"
/db_xref="taxon:195102"
/notes="anaerobic pathogen for gas gangrene"
64..342
/gene="CPE0728"
64..342
/gene="CPE0728"
/notes="92 aa, similar to pir:140869 hypothetical protein 4
nanH region from Clostridium perfringens (190 aa); 97.6%
identity in 84 aa overlap. Truncated by frameshift
mutation (confirmed by PCR-direct sequencing)"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB80434.1"
gene
CDS
/db_xref="GI:18144388"
/translation="MNKKVILITGASSGIGYSTAKLLLEBGIHVCTARRVNNRELO
RLGGKIIIFLDVKDERAVQVNTTIRMESRIDVVVFANSWGIRNCKYR"
290..790
/gene="CPE0729"
290..790
/gene="CPE0729"
/notes="166 aa, similar to pir:140869 hypothetical protein
4 nanH region from Clostridium perfringens (190 aa); 97.1%
identity in 103 aa overlap. Truncated by frameshift
mutation (confirmed by PCR-direct sequencing)"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB80435.1"
/db_xref="GI:18144389"
/translation="MWFLTLTGGALETNIDDAKEEDVNLFGVARVIAKAVMTQWRKQ
KYGIIVTSIAGIVSTPMSPFYPSKHALEGMIDGFRMENKESGIKIVKIQPFINT
EFIKPFMSNLKASKTKNAKAYKKEFEFFRKSSTKLILSSPPPEKVAKVSIIKKK
PKNSIL"
384..1448
/gene="CPE0730"
384..1448
/gene="CPE0730"
/notes="154 aa, similar to sp:YIAC.ECOLI HYPOTHETICAL 17.1
KDA PROTEIN IN TAG-BISC INTERGENIC REGION. from
Escherichia coli (146 aa); 36% identity in 139 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB80436.1"
/db_xref="GI:18144390"
/translation="MGERVLIRBPKNCEIDDIMKWLSTVEAHFYEEFYWKKNYE
WRDIVIPAKTEVYCDGSKINGPISIIIDNSFIGALFVHTKSGSGKSLKLVYVKNK
YANIELAVYKONKXAVEYKXKDFKILKEQENEDSGHLEYLMSYKSSSENQ"
1804..1980
/gene="CPE0731"
1804..1980
/gene="CPE0731"
/notes="58 aa, no significant homology."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB80437.1"
/db_xref="GI:18144391"
/translation="MKKCLYCKDLEKEPKESYIENKVGYFCSEDHFDKYILSTPEE
VIEVQNSFCVCSDD"
2233..2808
/gene="CPE0732"
2233..2808
/gene="CPE0732"
/notes="191 aa, similar to pir:H69339 conserved
hypothetical protein AF0720 from Archaeoglobus fulgidus
(214 aa); 38.5% identity in 213 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB80438.1"
/db_xref="GI:18144392"
/translation="MNEEDGKRSRAHSLFPSPVLKAAFAIGLPLFTAATNWDYKKNF
VKKLKEAKDLGAELVTGIDVDPEDCWHRVTNSIGLGLMPLMRKNHKEVEEFIN
LGFVTKIVTNLKNMKXEDLGRILTFDYIKELEERGIDPCGEAFETTTVIGGSLFK
KELRVKHGEIISGQNYMLTLELDFKLSD"
3023..3919
/gene="CPE0733"
3023..3919
/gene="CPE0733"
/notes="298 aa, similar to sp:YC30.ODOSI HYPOTHETICAL
TRANSCRIPTIONAL REGULATOR YCF30 from Odontella sinensis
chloroplast (309 aa); 29% identity in 262 aa overlap
LysR family"
/codon_start=1
/transl_table=11

```

```

/product="probable transcriptional regulator"
/protein_id="BAB80439.1"
/db_xref="GI:18144393"
/translation="MNLYLRSFYTTVKCNISKAQKQLHLTPQVSNQIQLENDIN
FKLNRSTGVSLSAGGIIIEFAEMLSLIEDNLOKLDLKNLSKLIISCCSKSLGE
HVMPCSYTFKEIHYKTDISWEIDNTSLDKLTNHTNIGIIQKPSKDKNLPL
MTDNLVSGNRTSKSVNVELYKPLITREDGSANKLLVESLKKKNKINIDNLFV
LSLNSPQISKSVFSGQYSELPAISLNLHRSSELKIDINDITLFPQYVIALRKY
KLDYEQFVDFLTKKRCFCV"
4269..5969
/gene="naox"
/gene="naox"
/length="5969"
/notes="566 aa, similar to gnu:AE004394.5 NADH oxidase,
putative from Vibrio cholerae (567 aa); 43% identity in
530 aa overlap. Putative N-terminal signal sequence was
found by PSORT
CPE0734"
/codon_start=1
/transl_table=11
/product="NADH oxidase"
/protein_id="BAB80440.1"
/db_xref="GI:18144394"
/translation="MSKIVIVGVAGGASTAARLRLENNQIMPERGPHVSFNC
CLPYLSGKVENHEDLVLPTEKFYSQVRIADRVYEVYSINRAKEVTVNVSQGE
YTESDKLVSPGAKAIVPPIKGIEDVNVFTVRNVDIADKLSPIKONNSKNIIVIG
FGIVEAENLEAGYVTLTEAMDOIIMKPDYDMVOVLHKEMYDKGNLVIVGDKVSS
PESGIVLESKKNDSAVNAIGVAPETDLAREAGLEIGOTGAIKYNQVLTNDKDI
YAVGDVAVYNSLINSISRLPLAGPAQQAARAVADHHRSPRSNSGIVSGCVQIFHY
NGASTLGEQKQATLNSLNDVVRVIFQDKVGLMPGSEPLHFKLIEVPVGRLLGAG
AIGNAKDRVDITATLTKMGTVEDDKLVCVAPFGTAKDVVNAALVASKLLEG
TFQYVHEVSDLVESGACIIDVRESEYEAGHKGAKNPLSEIRDLPIPTDRPV
YLHCRSAQSRYNACLALQHLGFDNIYVNSGGFMGISFYEVYNDVQDREPITVYENF
"
6075..7361
/gene="CPE0735"
/gene="CPE0735"
/length="7361"
/notes="428 aa, partially similar to pir:D82968
hypothetical protein PA5430 from Pseudomonas aeruginosa
(strain PA01) (404 aa); 25.5% identity in 314 aa overlap.
Putative N-terminal signal sequence and 7 putative
transmembrane regions were found by PSORT."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB80441.1"
/db_xref="GI:18144395"
/translation="WKXTEYIIGFILLAAIIIGKINLGSSMLPFRLLAGLGLYAL
TRSFPGAGVNRSTKSTKLMRLMLFVGTAVISVCFFIGQDPREYDLWNPIL
GLIIGLFGFMFSSCCASGVLTDVVTGLPRAFITLIPFGMGVYLFQVQNSAVW
KDTWISTFNGVFLPDLFKWDLGELVGLAIIVITILACTIIVAKLYBEKRSANNV
YTGVDTEILOQKHKEEKEFKLFSEATVEKLFVRPWTLATGALVIGGIFILLMGVTKA
GWGASTPYGFRALVAVGVSPEVSATFAGKVPVMTAPFSSNGMVQVNGIILGL
VALLAGSKFTFKAGLQISEKALLIYVGGFANGFGTSLNGCNVCAITYPIANLSL
SGWFLPLFVAGVGLGNIVYKIMGNKCKKEVINNK"
7450..8781
/gene="CPE0736"
/gene="CPE0736"
/length="8781"
/notes="443 aa, similar to gnu:AP001510.169 BH0994 gene
product from Bacillus halodurans (395 aa); 24.3% identity
in 369 aa overlap. Putative N-terminal signal sequence and
11 putative transmembrane regions were found by PSORT."
/codon_start=1
/transl_table=11
/product="probable transporter"
/protein_id="BAB80442.1"
/db_xref="GI:18144396"
/translation="WKEIKLSRIDFSVLGAILGWGSFMLPGTKFLKESGVNTFPG
LFLGVCIIIRIKYLLINMQTHDEGFSFTYNNLKGKHGFIIVGFWLTLAYFTMPL
NATAPFLVIFGGLIEFGYVNAVYNIYGLIELTSSIIIIAFALNINLNGIKTKS
VQNLIIFSLIGMTFVLVGMILKGNRIDFINTYINKYSFDSLQIKVFAITPPAFIFG

```

```

DAIPQISKELNFSKKASRVAVISLVGALIVNIIITALAYSPEQASLEWAQSA
VLSLKGAFPLLIILAIATAVWSGINGFMICSKLGSIANYKMLSRKMGKNGVIF
SNAIITIVLSIAPWGRQAIWIVMSSLSGASVAYFYVSVILKESNTKDKILAG
IGVIGIIFMLLIILIPSAALSKESLIALIWCIIIGFIAYVKIQDNIEA"
/length="8940"
/gene="CPE0737"
/length="8940"
/notes="8940 aa, similar to >gp:LM0132543.1
/notes="220 aa, similar to >gp:LM0132543.1
Query Match 18.4%; Score 40; DB 1; Length 301450;
Best Local Similarity 52.4%; Pred. No. 3;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
Qy 2 ATTCCGACATGATGATACAGCGGATTAATACAGGTAAGTTCATTACACGGAACCC 61
Db 214171 ATTATATATATTTTACTATAGATTATTAATAAAGGTGAATTTATGAAGATAGAAGC 214230
Qy 62 TTTAAGAATAATATGAAGGTATTACCTTGTGTCACCGGAATGGTAAATATGCCCG 121
Db 214231 TTCAAGTACCATAGTAGAGTTCAAGAGTATTTCACAGTTCATATTAATCAAGGG 214290
Qy 122 AGTTTTCATGATACGGAATCCAGCCATTTCTTATTCATATTTGACT 169
Db 214291 AGTTATAGACTATAAACTTCCTTTATGTGATTTAATGAGTAGTATCT 214338
RESULT 14
AF412737/c AF412737 695 bp DNA linear INV 26-SEP-2001
LOCUS Inachis io NADH dehydrogenase 1 (ND1) gene, partial cds;
DEFINITION Mitochondrial gene for mitochondrial product.
ACCESSION AF412737
VERSION AF412737.1 GI:15778377
KEYWORDS Mitochondrion Inachis io
SOURCE Inachis io
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Papilionoidea; Nymphalidae; Nymphalinae; Inachis.
REFERENCE 1 (bases 1 to 695)
AUTHORS Nylin,S., Nyblom,K., Ronquist,F., Janz,N., Belicek,J. and
Kallersjo,M.
TITLE Phylogeny of Polytonia, Nymphalis and related butterflies
(Lepidoptera: Nymphalidae): a total-evidence analysis
JOURNAL Zool. J. Linn. Soc. 132, 441-468 (2001)
REFERENCE 2 (bases 1 to 695)
AUTHORS Nylin,S., Nyblom,K., Ronquist,F., Janz,N., Belicek,J. and
Kallersjo,M.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-2001) Department of Zoology, Stockholm
University, Stockholm 106 91, Sweden
FEATURES Location/Qualifiers
source
1..695
/organism="Inachis io"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:171585"
<1..>695
/gene="ND1"
<1..>695
/gene="ND1"
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase 1"
/protein_id="AA07408.1"
/db_xref="GI:15778378"
/translation="MIDMENIMIGFLIIVGMIGVAFILIERKVLGIQIKGPNK
MGWGLLQPCDAIKLFSEQVLYNYSVLYFIPDINGFILSLXIMVYPMYFNMX
FNGLIFFXXCTQGVLYTMGMSSNSYSLGLGRAVQTIYSVEVMTIMMSSI
IIMDFNLKMFNYQVLYTMGMSSNSYSLGLGRAVQTIYSVEVMTIMMSSI
IEYSSGCGALIFLA"
BASE COUNT 216 a 54 c 99 g 316 t 10 others

```

```

ORIGIN
  Query Match      18.1%; Score 39.2; DB 3; Length 695;
  Best Local Similarity 56.1%; Pred. No. 13;
  Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
  QY 1 AATCGGACGATGATACAGCGGATTAATACAGGTAAGTTCAITACACGGAAGAAC 60
  DB 217 AATTGAATAATTTAAATAAACTTGTTCTTTAGAAAATAACTTAATTCGATCGAAAAAG 158
  QY 61 CTTTAAAGAAATATGAGGATTAACCTTGTTGCGAAGGGAATGGTAATATGCCCC 120
  DB 157 GTTGTAATACCCATATATAACCTATTTTATTTGGGCGCTTTCCGAATTGTAATACCTA 98
  QY 121 GAGTTTTCAC 132
  DB 97 AAACCTTTCGCT 86

RESULT 15
BX088572/c
LOCUS
DEFINITION
  Danio rerio clone CH211-246K16, *** SEQUENCING IN PROGRESS ***.
ACCESSION
  BX088572
VERSION
  BX088572.3 GI:29538729
KEYWORDS
  HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_PULLTOP.
SOURCE
  Danio rerio (zebrafish)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
  Cypriniformes; Cyprinidae; Danio.
  1 (bases 1 to 163453)
  McLaren,S.
  Direct Submission
  Submitted (03-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Apr 3, 2003 this sequence version replaced gi:28460219.
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: zfish-help@sanger.ac.uk
  ----- Project Information
  Center project name: zC246K16
  ----- Summary Statistics
  Assembly program: XGAP4; version 4.5
  Chemistry: Dye-terminator; 100% of reads
  Consensus quality: 163374 bases at least Q40
  Consensus quality: 163411 bases at least Q30
  Consensus quality: 163430 bases at least Q20
  Insert size: 163453; sum-of-contigs
  Insert size: 167545; 4.9% error; agarose-fp
  Quality coverage: 7.07x in Q20 bases; sum-of-contigs Quality
  coverage: 7.11x in Q20 bases; agarose-fp
  -----
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * been provided by the submitter.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  * 1 163453: contig of 163453 bp in length.
  Location/Qualifiers
  1..163453
   /organism="Danio rerio"
   /mol_type="genomic DNA"
   /db_xref="taxon:7955"
   /clone="CH211-246K16"
   /clone_lib="CHORI-211"
  
```

```

misc_feature 1..163453
  /note="assembly fragment:00934"
BASE COUNT 50337 a 31066 c 30938 g 51111 t 1 others
ORIGIN
  Query Match      18.1%; Score 39.2; DB 2; Length 163453;
  Best Local Similarity 58.6%; Pred. No. 5.2;
  Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
  QY 22 GCGGATTAATACAAAGTAAAGTTCAITACACGGAAGAACCTTTAAAGAAATATGAAAG 81
  DB 129808 GCAGATTCAGCAAGTCAAAATTAAGACTATTTAAGACCTTTTAAAGATTTATGATG 129749
  QY 82 GTATTACCTTGTTCCTCCCAACGGAATGTAATATGCCGAGTTTTTCACCTGAATA 137
  DB 129748 ATATCTTTTATATACAAAGGCGGAATGCTTAGAATGGACATGATTTTAAATGTAAA 129693

Search completed: November 15, 2003, 03:34:21
Job time : 965.589 secs
  
```

FEATURES
source

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:55:19 ; Search time 140.052 Seconds
(without alignments)
4182.570 Million cell updates/sec

Title: US-09-928-457-99

Perfect score: 217

Sequence: 1 RAATCGGACAGTATCAATAC.....ATGACGATGAAGATTTAATT 217

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	100.0	217	AAV30599	Neisseria meningitidis
2	130.4	60.1	858	AA15296	DNA encoding a pol
3	130.4	60.1	858	AA15322	DNA encoding a pol
4	128.4	59.2	855	AB237755	N. gonorrhoeae nuc
5	128.4	59.2	891	AB241353	N. gonorrhoeae nuc
6	37.6	17.3	472	AB058421	Human colon cancer
7	37.6	17.3	602	ABK55260	Human colon cancer
8	37.6	17.3	2998	AA182276	Lung cancer associ

C 9	37.6	17.3	3817	22	AAH02874	Human shear stress
C 10	37.6	17.3	3826	24	ABX84707	Human cDNA differe
C 11	36	16.6	2000	24	ABZ15907	Arabidopsis thalia
C 12	36	16.6	640681	24	ABX92787	Buchnera sp. genom
C 13	35.4	16.3	163350	24	AA46127	Human tumour suppr
C 14	34.8	16.0	17421	22	AA545349	Chemically pretrea
C 15	34.8	16.0	17421	24	ABX28182	DNA transcription
C 16	34.4	15.9	339	22	ABAI2692	Human nervous syst
C 17	34.4	15.9	393	22	ABA18110	Human nervous syst
C 18	34.4	15.9	393	22	ABA18111	Human nervous syst
C 19	34.4	15.9	393	22	ABA18112	Human nervous syst
C 20	34.2	15.8	2047	17	AA114767	Campylobacter jeju
C 21	33.8	15.6	584	23	ABV52653	Human prostate exp
C 22	33.6	15.5	540	24	ABX60857	Human cancer relat
C 23	33.6	15.5	4873	22	AAK81633	Human immune/haema
C 24	33.2	15.3	1498	23	ABV20826	Human prostate exp
C 25	33.2	15.3	1498	23	ABV21352	Human prostate exp
C 26	33.2	15.3	1498	23	ABV21414	Human prostate exp
C 27	33.2	15.3	1498	23	ABV21491	Human prostate exp
C 28	33.2	15.3	1498	23	ABV21495	Human prostate exp
C 29	33.2	15.3	1498	23	ABV21507	Human prostate exp
C 30	33.2	15.3	1498	23	ABV21565	Human prostate exp
C 31	33.2	15.3	1498	23	ABV21675	Human prostate exp
C 32	33.2	15.3	1498	23	ABV21700	Human prostate exp
C 33	33.2	15.3	1498	23	ABV26673	Human prostate exp
C 34	33.2	15.3	1498	23	ABV27171	Human prostate exp
C 35	33.2	15.3	1498	23	ABV27232	Human prostate exp
C 36	33.2	15.3	1498	23	ABV27309	Human prostate exp
C 37	33.2	15.3	1498	23	ABV27313	Human prostate exp
C 38	33.2	15.3	1498	23	ABV27325	Human prostate exp
C 39	33.2	15.3	1498	23	ABV27384	Human prostate exp
C 40	33.2	15.3	1498	23	ABV27496	Human prostate exp
C 41	33.2	15.3	1498	23	ABV27522	Human prostate exp
C 42	33	15.2	334	19	AAV03336	DNA sequence that
C 43	32.8	15.1	18434	24	ABJ34006	Human immune syste
C 44	32.6	15.0	301	25	ABD53229	Human chromosome 3
C 45	32.6	15.0	461	22	AAH13064	Human cDNA clone (

ALIGNMENTS

RESULT 1
AAV30599
ID AAV30599 standard; DNA; 217 BP.
XX
AC AAV30599;
XX
DT 22-OCT-1998 (first entry)
XX
DE Neisseria meningitidis DNA sequence E110.
XX
KW N. Gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW meningitis; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9802547-A2.
XX
PD 22-JAN-1998.
XX
PF 11-JUL-1997; 97WO-FR01295.
XX
PR 12-JUL-1996; 96FR-0008768.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (PLAC) MAX BLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (SMIX) SMITHKLINE BEECHAM.
XX
PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI Vinals C;
XX

DR WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX

PS Example 4; Page 131; 150pp; French.

XX

XX This sequence represents a sequence present in *Neisseria meningitidis*

CC and *N. gonorrhoeae* but not in *N. lactamica*, except for the genes involved

CC in biosynthesis of the capsule polysaccharide, *frpA* or *C*, *opc*, *porA*,

CC rotamase, sequence IC1106, IGA protease, pillin, *pilC*, proteins which

CC bind transferrin and opacity proteins. The DNA sequences are responsible

CC for the differences in pathogenicity between *N. meningitidis* and

CC *N. gonorrhoeae*, specifically they include the genes that allow

CC *N. meningitidis* to cross the blood-brain barrier. DNA sequences common

CC to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N. lactamica*,

CC are responsible for colonisation and penetration of the mucosa. The

CC DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX

SQ Sequence 217 BP; 76 A; 31 C; 44 G; 66 T; 0 other;

XX

Query Match 100.0%; Score 217; DB 19; Length 217;

Best Local Similarity 100.0%; Pred. No. 6.2e-53;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGATACAGCGGATTAATACAGGTAGTTTCAATACACGGAAAAAC 60

DB 1 AATTCGGACAGTATGATACAGCGGATTAATACAGGTAGTTTCAATACACGGAAAAAC 60

QY 61 CTTTAAAGAAATATGAAGGTATTACCTTTGTTGCCAACGGGAATGTTAAATATGCC 120

DB 61 CTTTAAAGAAATATGAAGGTATTACCTTTGTTGCCAACGGGAATGTTAAATATGCC 120

QY 121 GAGTTTTTCATGTAATACGGAATCCAGCCATTTCTATTCATATTGACTGGATGGCTGAA 180

DB 121 GAGTTTTTCATGTAATACGGAATCCAGCCATTTCTATTCATATTGACTGGATGGCTGAA 180

QY 181 TGTGCACATTTAGATATGATGACGATGAGATTTAATT 217

DB 181 TGTGCACATTTAGATATGATGACGATGAGATTTAATT 217

RESULT 2

AAAL5296

ID AAAAL5296 standard; DNA; 858 BP.

XX

XX AAAL5296;

XX

DT 04-SEP-2000 (first entry)

DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.

XX

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

XX

XX *Neisseria meningitidis*.

XX

XX Key Location/Qualifiers

PH Key 1..858

FT CDS /*tag= a

FT

XX WO200026375-A2.

XX

XX 11-MAY-2000.

XX

XX 28-OCT-1999; 99WO-FR02643.

XX

XX 30-OCT-1998; 98FR-0013693.

XX

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

XX Tinsley C, Perrin A;

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX

XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

XX Tinsley C, Perrin A;

XX

XX WPI; 2000-365622/31.

DR P-PSDB; AAY93265.

XX

XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic

PT or preventative vaccines and for diagnosis

XX

XX Claim 2; Page 51-52; 187pp; French.

XX

XX The present sequence encodes a protein that is specific for pathogenic

CC strains of *Neisseria*. The polynucleotides, polypeptides, or their

CC antigenic fragments, are used in vaccines to treat or protect against

CC *Neisseria* infections, particularly by *N. meningitidis*. The

CC polynucleotide sequence is also used for recombinant production of

CC the polypeptide and to produce attenuated *Neisseria* strains that

CC overexpress it, or express it in a non-toxic mutant form.

XX

SQ Sequence 858 BP; 297 A; 145 C; 182 G; 234 T; 0 other;

XX

Query Match 60.1%; Score 130.4; DB 21; Length 858;

Best Local Similarity 97.5%; Pred. No. 6.7e-28;

Matches 154; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 AATTCGGACAGTATGATACAGCGGATTAATACAGGTAGTTTCAATACACGGAAAAAC 60

DB 701 AATTCGGACAGTATGATACAGCGGATTAATACAGGTAGTTTCAATACACGGAAAAAC 760

QY 61 CTTTAAAGAAATATGAAGGTATTACCTTTGTTGCCAACGG--GAATGGTAAATATGC 118

DB 761 CTTTAAAGAAATATGAAGGTATTACCTTTGTTGCCAACGGAGATGTTAAATATGC 820

QY 119 CCGAGTTTTTCACTGTAATAGCGAATCCAGCCATTTCTA 156

DB 821 CCGAGTTTTTCACTG-ATAGCGAATCCAGCCATTTCTA 857

RESULT 3

AAAL5322

ID AAAAL5322 standard; DNA; 858 BP.

XX

XX AAAL5322;

XX

DT 04-SEP-2000 (first entry)

DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.

XX

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

XX

XX *Neisseria gonorrhoeae*.

XX

XX Key Location/Qualifiers

PH Key 1..858

FT CDS /*tag= a

FT

XX WO200026375-A2.

XX

XX 11-MAY-2000.

XX

XX 28-OCT-1999; 99WO-FR02643.

XX

XX 30-OCT-1998; 98FR-0013693.

XX

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

XX Tinsley C, Perrin A;


```
DR WPI; 2000-365622/31.
DR P-PSDB; AAY93292.
XX
PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT or preventative vaccines and for diagnosis -
XX
XX PS Claim 4; Page 134-135; 187pp; French.
XX
CC The present sequence encodes a protein that is specific for pathogenic
CC strains of Neisseria. The polynucleotides, polypeptides, or their
CC antigenic fragments, are used in vaccines to treat or protect against
CC Neisseria infections, particularly by N. meningitidis. The
CC polynucleotide sequence is also used for recombinant production of
CC the polypeptide and to produce attenuated Neisseria strains that
CC overexpress it, or express it in a non-toxic mutant form.
XX
XX SQ Sequence 858 BP; 302 A; 151 C; 178 G; 227 T; 0 other;

Query Match 60.1%; Score 130.4; DB 21; Length 858;
Best Local Similarity 97.5%; Pred. No. 6.7e-28;
Matches 154; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 AATTGGACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGGAARAC 60
Db 701 AATTGCAACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGGAARAC 760

QY 61 CTTTAAAGAATAATATGAAGAAGTATTACCTTTGTTGCCAACGG--GAATGTTAAATATGC 118
Db 761 CTTTAAAGAATAATATGAAGAAGTATTACCTTTGTTGCCAACGGGAATGTTAAATATGC 820

QY 119 CCGAGTTTTTCACTGAATAGCGAATCCAGCCATTCTA 156
Db 821 CCGAGTTTTTCACTG-ATAGCGAATCCAGCCATTCTA 857

RESULT 4
ABZ37755
ID ABZ37755 standard; DNA; 855 BP.
XX
XX AC ABZ37755;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 99.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
PN WPI; 2003-058415/05.
DR P-PSDB; ABP76785.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
XX
XX PS Claim 6; Page 188; 815pp; English.
XX
CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention.
XX
XX SQ Sequence 891 BP; 316 A; 153 C; 186 G; 236 T; 0 other;

Query Match 59.2%; Score 128.4; DB 25; Length 891;
Best Local Similarity 97.4%; Pred. No. 2.6e-27;
Matches 152; Conservative 0; Mismatches 1; Indels 3; Gaps 2;
```

```
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention.
XX
XX SQ Sequence 855 BP; 300 A; 151 C; 178 G; 226 T; 0 other;

Query Match 59.2%; Score 128.4; DB 25; Length 855;
Best Local Similarity 97.4%; Pred. No. 2.5e-27;
Matches 152; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 AATTGGACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGGAARAC 60
Db 701 AATTGCAACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGGAARAC 760

QY 61 CTTTAAAGAATAATATGAAGAAGTATTACCTTTGTTGCCAACGG--GAATGTTAAATATGC 118
Db 761 CTTTAAAGAATAATATGAAGAAGTATTACCTTTGTTGCCAACGGGAATGTTAAATATGC 820

QY 119 CCGAGTTTTTCACTGAATAGCGAATCCAGCCATTTC 154
Db 821 CCGAGTTTTTCACTG-ATAGCGAATCCAGCCATTTC 855

RESULT 5
ABZ41353
ID ABZ41353 standard; DNA; 891 BP.
XX
XX AC ABZ41353;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 7295.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
PN WPI; 2003-058415/05.
DR P-PSDB; ABP80383.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
XX
XX PS Disclosure; Page 718; 815pp; English.
XX
CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention.
XX
XX SQ Sequence 891 BP; 316 A; 153 C; 186 G; 236 T; 0 other;

Query Match 59.2%; Score 128.4; DB 25; Length 891;
Best Local Similarity 97.4%; Pred. No. 2.6e-27;
Matches 152; Conservative 0; Mismatches 1; Indels 3; Gaps 2;
```


Db 357 ACTTCTGCAAAATTCACAGGCACAGTGTTCATTCAACAGAGAAAAGTCAAACCACT 298
 QY 121 GAGTTTTCACGAATAGGAATCCAGCCATTTCTATTATATTT 164
 Db 297 TGGTTTTTAATGAAATCCTTCACATCCACCTGTGTTCAAAAT 254

RESULT 8
 AAF18276/c
 ID AAF18276 standard; DNA; 2998 BP.
 XX
 AC AAF18276;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polynucleotide sequence SEQ ID 295.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200055180-A2.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 XX Ruben SM;
 XX
 DR WPI; 2000-587514/55.
 DR P-PSDB; AAB58400.
 XX

Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 1; Page 754-755; 1425pp; English.

Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective, cytostatic, cardioactive, and
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.

Sequence 2998 BP; 826 A; 568 C; 699 G; 906 T; 9 other;
 SQ

Query Match 17.3%; Score 37.6; DB 21; Length 2998;
 Best Local Similarity 51.8%; Pred. No. 0.43;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACAGCGGATTAATCAAGGTAAAGTTCATTACACGGAATAAC 60
 Db 2282 AACTGAGGCATTAAATACCTCTTTTATAATATGCAAGTTGAAATGCTAACAAAGCATAAAC 2223
 QY 61 CTTTAAAGATAATATGAAAGGTATTACCTTGTGGCCAAACGGGAATGTAATAATGCCC 120
 Db 2222 ACTTCTGCAAAATTCACAGGCACAGTGTTCATTCAACAGAAAAGTCAAAACCACT 2163
 QY 121 GAGTTTTCACGTGAATAGCGAATCCAGCCATTTCTATTATATTT 164
 Db 2162 TGGTTTTTAATGAAATCCTTCACATCCACCTGTGTTCAAAAT 2119

RESULT 9
 AAH02874/c
 ID AAH02874 standard; DNA; 3817 BP.
 XX
 AC AAH02874;
 XX
 DT 15-JUN-2001 (first entry)
 XX
 DE Human shear stress-response coding sequence SEQ ID NO: 1.
 XX
 KW Human; shear stress-response protein; vascular disease;
 KW arteriosclerosis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200125427-A1.
 XX
 PD 12-APR-2001.
 XX
 XX 02-OCT-2000; 2000WO-JP06840.
 XX
 XX 01-OCT-1999; 99JP-0280976.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA (NOJI) NOJIMA H.
 XX
 XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 PI
 DR WPI; 2001-266308/27.
 DR P-PSDB; AAB90751.
 XX

DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis -
 XX
 PS Claim 20; Page 144-149; 678pp; Japanese.

The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.

Sequence 3817 BP; 1051 A; 736 C; 936 G; 1094 T; 0 other;
 SQ

Query Match 17.3%; Score 37.6; DB 22; Length 3817;
 Best Local Similarity 51.8%; Pred. No. 0.46;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACAGCGGATTAATCAAGGTAAAGTTCATTACACGGAATAAC 60
 Db 3162 AACTGAGGCATTAAATACCTCTTTTATAATGCAAGTTGAAATGCTAACAAAGCATAAAC 3103
 QY 61 CTTTAAAGATAATATGAAAGGTATTACCTTGTGGCCAAACGGGAATGTAATAATGCCC 120
 Db 3102 ACTTCTGCAAAATTCACAGGCACAGTGTTCATTCAACAGAAAAGTCAAAACCACT 3043
 QY 121 GAGTTTTCACGTGAATAGCGAATCCAGCCATTTCTATTATATTT 164

DB 3042 TGGTTTAAATGAATCCTTCACATCCACCTGTTGTTCAAAAT 2999

RESULT 10
ID ABK84707/c
XX ABK84707 standard; cDNA; 3826 BP.
AC ABK84707;
XX
XX 14-AUG-2002 (first entry)
XX Human cDNA differentially expressed in granulocytic cells #1278.
DE
XX
XX Human; ss: granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
XX Homo sapiens.
XX
XX WO200228999-A2.
XX
XX 11-APR-2002.
XX
XX 03-OCT-2001; 2001WO-US30821.
XX
XX 03-OCT-2000; 2000US-237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity
XX
XX Claim 1; SEQ ID No 1278; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GCA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,

CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3826 BP; 1053 A; 740 C; 937 G; 1096 T; 0 other;
XX
XX Query Match 17.3%; Score 37.6; DB 24; Length 3826;
XX Best Local Similarity 51.8%; Pred. NO. 0.46;
XX Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
QY 1 AATTCGGACAGTATCAATACAGCGGATTAAATACAGGTAAGTTCATTACACGGAAGAAC 60
DB 3162 AACTGAGGCAATTAACCTCTTTATTAATGCAAGTTGAATGCTAACAGCATAAAC 3103
QY 61 CTTTAAAGAATAATGAAAGTATTACCTTGTTCGCAACGGAATGTTAAATATGCC 120
DB 3102 ACTTCTGCAAAAATTCACAAAGGCACAGTTGTTTCATTCACAGAAAAAGTCAAAACCACT 3043
QY 121 GAGTTTTCACGTAGTACGAATCCAGCCATTTCTATTCAATATT 164
DB 3042 TGGTTTAAATGAATCCTTCACATCCACCTGTTGTTCAAAAT 2999

RESULT 11
ABZ15907
ID ABZ15907 standard; DNA; 2000 BP.
XX
XX AC ABZ15907;
XX
XX 21-JAN-2003 (first entry)
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 3712.
DE
XX Arabidopsis thaliana, plant; gene; stress; transgenic; ds.
XX
XX Arabidopsis thaliana.
OS
XX Arabidopsis thaliana.
XX
XX WO200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26685.
XX
XX 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-264647P.
PR 22-JUN-2001; 2001US-300111P.
XX
XX (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX Claim 144; SEQ ID NO 3712; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX Sequence 2000 BP; 768 A; 336 C; 274 G; 622 T; 0 other;

Query Match 16.6%; Score 36; DB 24; Length 2000;
Best Local Similarity 51.9%; Pred. No. 1.1;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 29 AATCAAGTAAGTTCATTACACGGAACCCCTTAAAGATAATATGAAGGTATTAC 88
DB 13 AAACATCGACGAGAAATTTAATACGTAAGCACCATAAAATGAGAAAAATACACAACTAT 72
OY 89 CTGTGTTGCCACGGGAATGTAATATGCGCGAGTTTTCACCTCAATAGCGAATCCAGC 148
DB 73 AGTGATGACGATGCTTTTGACATATAAGTCTTCTCGTTAACAGATTATAATCTATT 132
OY 149 CATTCTATTTCATATTGACTGGATGGCTGAATGTG 184
DB 133 TTTTTCGATTGATCTTACGCTTATGCAAAATTATG 168

RESULT 12
ABA92787
ID ABA92787 standard; DNA; 640681 BP.

XX ABA92787;

DT 27-MAR-2002 (first entry)

XX Buchnera sp. genomic DNA SEQ ID NO:1.

DE Buchnera; cockroach-symbiotic bacterium; cockroach extermination;
KW circular; ds.

XX Buchnera sp.

OS JP2001292771-A.

PN 23-OCT-2001.

XX 07-APR-2000; 2000JP-0107160.

XX 07-APR-2000; 2000JP-0107160.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 2002-126043/17.

XX A genomic DNA of cockroach-symbiotic bacterium -

PS Claim 1; Page 16-230; 237pp; Japanese.

CC The present invention describes a Gene (I) derived from Buchnera sp.
CC containing the DNA (a) or (b). (a) has a fully defined base pair
CC sequence selected from a table of sequences found in the Buchnera sp.
CC genomic DNA of ABA92787 given in the specification or is a DNA selected
CC from complementary DNA sequences, and (b) is a DNA which hybridises with
CC the DNA (a) and encodes a protein. Also described are: (1) a recombinant
CC vector (ii) containing (I); (2) a transformant (iii) containing (II);
CC (3) a genomic DNA of Buchnera sp. containing the sequence given in
CC ABA92787; (4) a plasmid derived from Buchnera sp. containing DNA (c) or
CC (d). (c) is a DNA containing a fully defined sequence given in ABA92788
CC or ABA92789 and (d) is a plasmid which hybridises with a DNA; and (5) a
CC method for the preparation of a protein in which (iii) is cultured and
CC the expression protein of the objective protein is collected from the
CC resultant culture. The DNA is useful for developing agricultural
CC chemicals for exterminating cockroaches. The present sequence represents
CC the specifically claimed Buchnera sp. genomic DNA sequence, from the

CC present invention.

XX Sequence 640681 BP; 237522 A; 83822 C; 84757 G; 234580 T; 0 other;

Query Match 16.6%; Score 36; DB 24; Length 640681;
Best Local Similarity 55.6%; Pred. No. 4.2;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 27 TTAATCAAGTAAGTTCATTACACGGAACCCCTTAAAGATAATATGAAGGTATT 86
DB 512832 TTAATTAAGCATCCAAAGAAATAGGAAATCTAAATGAATAGTATTATTTATTT 512891
OY 87 ACCTTGTTCCCAACGGGAATGTAATATATGCGCGAGTTTTCACCTGAATAGCGAATCCA 146
DB 512892 AAACCTCTTCTTAAGAGACCTGCTAAACAGCCAGCAAAATTCCTGGAATACCTAACCA 512951
OY 147 GCCA 150
DB 512952 GTCA 512955

RESULT 13
AAD46127/c
ID AAD46127 standard; DNA; 163350 BP.

XX AAD46127;

DT 27-DEC-2002 (first entry)

XX Human tumour suppressor gene.

DE Human; tumour suppressor protein; cell proliferative disorder; vaccine;
KW inflammation; brain cancer; adenocarcinoma; cervix cancer; bone cancer;
KW apoptosis; leukaemia; lymphoma; melanoma; therapy; chromosome 13; gene;
KW ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2001..160511

FT /*tag= a

FT /product= "Human tumour suppressor protein"

FT exon 2001..2178

FT /*tag= b

FT intron 2179..159807

FT /*tag= c

FT exon 159808..160511

FT /*tag= d

XX WO200268468-A2.

XX 06-SEP-2002.

XX 05-FEB-2002; 2002WO-US03235.

XX 27-FEB-2001; 2001US-0793706.

XX (PEKE) PE CORP NY.

XX Gong F, Yan C;

XX WPI; 2002-698658/75.

XX P-PSDB; AAB28633.

XX New human tumor suppressor proteins, useful for developing human
XX therapeutic agents, or preventing or treating inflammation, or
XX disorders associated with cell proliferation, e.g. bone cancer, brain
XX cancer, leukemia or lymphoma -

XX Claim 4; Fig 3; 200pp; English.

XX The invention relates to human tumour suppressor polypeptides and
XX polynucleotides. Sequences of the invention are useful for diagnosing,

CC preventing or treating inflammation, or disorders associated with cell
 CC proliferation and apoptosis e.g. bone cancer, brain cancer, cervix
 CC cancer, adenocarcinoma, leukaemia, lymphoma or melanoma. They are
 CC particularly useful as models for developing human therapeutic targets,
 CC identifying therapeutic proteins, or serving as targets for the
 CC development of human therapeutic agents that modulate tumour suppressor
 CC protein activity in cells and tissues that express the tumour suppressor
 CC protein. Polypeptides of the invention are used for identifying agents
 CC that modulate their activity. They are useful for raising antibodies or
 CC eliciting an immune response; as a reagent in assays designed to
 CC quantitatively determine levels of the protein (or its binding partner
 CC or ligand) in biological fluids; or as markers for tissues in which the
 CC corresponding protein is preferentially expressed. The invention is also
 CC used as vaccines. The present sequence is human tumour suppressor gene
 CC located on chromosome 13.

SQ Sequence 163350 BP; 40430 A; 36356 C; 37002 G; 39020 T; 10542 other;

Query Match 16.3%; Score 35.4; DB 24; Length 163350;
 Best Local Similarity 49.7%; Pred No. 4.6;
 Matches 90; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 37 GTAAGTTCATTACACCGGAAACCTTTAAGAAATATATGAAAGGTATTACCTGTGTTG 96
 Db 163047 GCAACTGCTATTCATTGAATCTTCATTTCTATAACTTAATTGCAATTATACCTCCAATT 162988

QY 97 CCAACGGGAATGGTAATATGCCGAGTTTTTCACTGAATAGCAATCCAGCCATTTCCTA 156
 Db 162987 AATATAGAAAGTGAAATATTAGCATGAATTTTATATAGCAAGACATCAAGCTATTTTA 162928

QY 157 TTCATATTGTACTGATCGCTGAATGTGGACTTTTATAGATAATGACCATGAAGATTTAAT 216
 Db 162927 TTGTCACCTGGATGTTTATTAGTATTTTAGTTTCATTAACCTTTTGT 162868

QY 217 T 217
 Db 162867 T 162867

RESULT 14
 AAS45349/c
 ID AAS45349 standard; DNA; 17421 BP.
 XX
 AC AAS45349;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Chemically pretreated complementary DNA associated with cell cycle #27.
 XX
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 PN WO200168911-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP02945.
 XX
 PR 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602751/68.
 DR
 XX Designing primers and probes for analysing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle -
 XX
 PS Claim 1; SEQ ID No 54; 28pp; English.
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers.
 XX
 SQ Sequence 17421 BP; 4578 A; 328 C; 3973 G; 8502 T; 40 other;

Query Match 16.0%; Score 34.8; DB 22; Length 17421;
 Best Local Similarity 65.4%; Pred No. 4.1;
 Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 AATACAGCGGATTAATACAGTAAAGTTCATTACACGGAACACCTTTAAGATATA 75
 Db 6042 AATATAAATATTAATACCATATAATTTTATTTTACCATATAAATAAATAATA 5983

QY 76 TGAAGGTTATACCTTGT 93
 Db 5982 TATAAATATATCTTTT 5965

RESULT 15
 ABK28182/c
 ID ABK28182 standard; DNA; 17421 BP.
 XX
 AC ABK28182;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA transcription associated complementary genomic DNA #28.
 XX
 KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KW immunological disorder; Werner syndrome; developmental disorder;
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
 KW polyglutamine disorder; solid tumour.
 XX
 OS Unidentified.
 XX
 PN WO200192565-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03973.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.
XX PA
XX PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-090046/12.
XX DR
XX PT New nucleic acids or oligomers, useful for diagnosing or treating
XX PT diseases associated with DNA transcription, e.g. immunological
XX PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid
XX PT tumours or cancer
XX PS
XX Claim 1; SEQ ID No 56; 32pp; English.
XX PS
XX The invention relates to a nucleic acid, which comprises a segment of the
XX CC chemically pretreated DNA of genes associated with DNA transcription from
XX CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
XX CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
XX CC to the chemically pretreated DNA of genes associated with DNA
XX CC transcription. The set of oligomer probes are useful for detecting the
XX CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
XX CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
XX CC diagnosing or treating diseases associated with DNA transcription
XX CC (particularly with the methylation status), e.g. adenosine deaminase
XX CC deficiency, viral infection, retroviral infection, Sezary syndrome,
XX CC haematological disorders, immunological disorders, Werner syndrome,
XX CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
XX CC neurological disorders, neurodegenerative disorders, Waardenburg
XX CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
XX CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
XX CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
XX CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
XX CC associated genomic DNA molecules of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from the
XX CC European Patent Office.
SQ Sequence 17421 BP; 4578 A; 328 C; 3973 G; 8502 T; 40 other;
Query Match 16.0%; Score 34.8; DB 24; Length 17421;
Best Local Similarity 65.4%; Pred No. 4.1;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 16 AATACAGCGGATTAATACAAGGTAAAGTTCATTACACGAAAAACCTTTAAAGATAATA 75
DB 6042 AATATAAAATATTAATACCATATATTTTATTACCATATAAAACATAAAAAATAATA 5983
QY 76 TGAAGGTATTACCTTGT 93
DB 5982 TATAAATTATATCTTTT 5965
Search completed: November 15, 2003, 00:35:51
Job time : 141.052 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 23:58:39 / Search time 1447.63 Seconds
(without alignments)
3643.257 Million cell updates/sec

Title: US-09-928-457-99

Perfect score: 217
Sequence: 1 AATTGGACAGTATGAATAC.....ATGACGATGAAGATTAAATT 217

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_esthum:*
4: em_esthum:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	39.2	18.1	445	12 BM796121	BM796121 K-EST0078
C 2	39.2	18.1	682	14 CB851528	CB851528 UI-CF-EN1
C 3	38.8	17.9	556	14 CA439981	CA439981 UI-H-DIO-
4	38.6	17.8	600	28 AZ525328	AZ525328 242PG08

5	38.6	17.8	956	28	BH146874	BH146874	ENT015TR
C 6	38.2	17.6	380	9	AA294909	AA294909	EST100072
C 7	37.6	17.3	308	9	AA356543	AA356543	EST65107
C 8	37.6	17.3	312	14	T32582	EST51185	Hu
C 9	37.6	17.3	324	9	AA294908	EST100071	
C 10	37.6	17.3	340	14	RS5154	Y44905.1	
C 11	37.6	17.3	402	12	B1017065	PM3-ET027	
C 12	37.6	17.3	407	10	BF446491	7Q88A03.X	
C 13	37.6	17.3	435	12	BM755155	K-EST0032	
C 14	37.6	17.3	437	14	N57411	Y482C07.1	
C 15	37.6	17.3	441	14	H48409	Y432F06.1	
C 16	37.6	17.3	489	10	BF565645	UI-R-BT1-	
C 17	37.6	17.3	511	14	CB120144	K-EST0167	
C 18	37.6	17.3	515	10	AW959021	EST371091	
C 19	37.6	17.3	518	12	BM888435	TM094	Hum
C 20	37.6	17.3	541	12	BM747615	K-EST0022	
C 21	37.6	17.3	541	12	BM744905	K-EST0018	
C 22	37.6	17.3	602	12	BM785453	K-EST0063	
C 23	37.6	17.3	620	12	BM743468	K-EST0016	
C 24	37.6	17.3	647	12	BM761014	K-EST0041	
C 25	37.6	17.3	675	12	BM706910	UI-E-COO-	
C 26	37.6	17.3	676	14	CB115165	K-EST0159	
C 27	37.6	17.3	682	9	AV682512	AV682512	
C 28	37.6	17.3	684	12	BM782216	K-EST0059	
C 29	37.6	17.3	754	10	BG576801	602599137	
C 30	37.6	17.3	775	14	CB963866	AGENCOURT	
C 31	37.6	17.3	782	10	BE545004	601070173	
C 32	37.6	17.3	819	10	BG168741	BG168741	602120021
C 33	37.6	17.3	830	13	BU933856	AGENCOURT	
C 34	37.6	17.3	832	12	B1260613	602968101	
C 35	37.6	17.3	848	10	BG171690	602322213	
C 36	37.6	17.3	864	13	BQ433011	AGENCOURT	
C 37	37.6	17.3	873	13	BQ953645	AGENCOURT	
C 38	37.6	17.3	880	13	BQ430231	AGENCOURT	
C 39	37.6	17.3	883	13	BQ435954	AGENCOURT	
C 40	37.6	17.3	884	13	BQ217267	AGENCOURT	
C 41	37.6	17.3	899	13	BU178416	AGENCOURT	
C 42	37.6	17.3	920	13	BX393445	EX393445	
C 43	37.6	17.3	923	13	BX393444	EX393444	
C 44	37.6	17.3	925	10	BE782607	601465726	
C 45	37.6	17.3	925	10	BE910626	601502609	

ALIGNMENTS

RESULT 1
BM796121/c
LOCUS
DEFINITION K-EST0078756 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-53-C04
5', mRNA sequence.

ACCESSION BM796121
VERSION BM796121.1
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 445)

AUTHORS Kim,N.S., Kahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.P., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,O.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 53 row: C column: 04
High quality sequence stop: 445.

```

FEATURES
  source
    Location/Qualifiers
      1..445
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="S22SNU16n1-53-C04"
        /sex="P"
        /tissue_type="Ascites"
        /cell_type="Lymphoblast-like"
        /cell_line="SNU-16"
        /lab_host="DH10B"
        /clone_lib="S22SNU16n1"
        /note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The S22SNU16 library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."
      BASE COUNT      108 a      75 c      90 g      172 t
      ORIGIN
        Query Match      18.1%; Score 39.2; DB 12; Length 445;
        Best Local Similarity 52.4%; Pred. No. 14;
        Matches 86; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
      QY 1 AATCGGCACATGATACAGCGGATTATACAGGTAAGTTCATTACACGGAAAAAC 60
      Db 236 AACTGAGGCATTAACTACCTTTTATAATATGCAAGTTGAAATGCTAACAAAGCATAAAC 177
      QY 61 CTTTAAAGATAATATGAAGGTATTACCTTGTTCACACGGGAATGGTAAATATGCC 120
      Db 176 ACTTCTGAAAATTCACAGGGCACAGTGTTCATTACACAGAAAAGTCAAACCCT 117
      QY 121 GAGTTTTCATGATACGGAATCCAGCCATTTCTATTCATATT 164
      Db 116 TGGTTTTAAATGAAATCCCTTCATCCACCTGTGTCAAAAT 73

RESULT 2
CB851528/c
LOCUS
DEFINITION
  UI-CF-EN1-aef-o-19-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
  UI-CF-EN1-aef-o-19-0-UI 3', mRNA sequence.
ACCESSION
  CB851528
VERSION
  CB851528.1 GI:30046303
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 682)
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  8889548
  Contact: McCray, PB
  McCray Lab
  2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
  Tel: 319 356 4866
  Fax: 319 356 7171
  Email: paul-mccray@uiowa.edu
  Tissue Procurement: Dr. M. J. Welsh, University of Iowa
  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com)

FEATURES
  source
    Location/Qualifiers
      1..682
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="UI-CF-EN1-aef-o-19-0-UI"
        /tissue_type="Primary Lung Cystic Fibrosis Epithelial
        Cells"
        /dev_stage="Adult"
        /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
        /clone_lib="UI-CF-EN1"
        /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
        modified polylinker; Site 1: EcoR I; Site 2: Not I;
        UI-CF-EN1 is a normalized cDNA library containing the
        following tissue(s): Primary Lung Cystic Fibrosis
        Epithelial Cells. The library was constructed according to
        Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
        1996. First strand cDNA synthesis was primed with an
        oligo-dT primer containing a Not I site. Double stranded
        cDNA was ligated to an EcoR I adaptor, digested with Not
        I, and cloned directionally into pT73-Pac vector. The
        oligonucleotide used to prime the synthesis of
        first-strand cDNA contains a library tag sequence that is
        located between the Not I site and the (dr)18 tail. The
        sequence tag for this library is CTGCTCAGGT.
        TAG SEQ=None found"
      BASE COUNT      189 a      104 c      138 g      251 t
      ORIGIN
        Query Match      18.1%; Score 39.2; DB 14; Length 682;
        Best Local Similarity 52.4%; Pred. No. 15;
        Matches 86; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
      QY 1 AATTCGGACAGTATGAATACAGCGGATTATACAGGTAAGTTCATTACACGGAAAAAC 60
      Db 657 AACTGAGGCATTAACTACCTTTAGTAATATGCAAGTTGAAATGCTAACAAAGCATAAAC 598
      QY 61 CTTTAAAGATAATATGAAGGTATTACCTTGTTCACACGGGAATGGTAAATATGCC 120
      Db 597 ACTTCTGAAAATTCACAGGGCACAGTGTTCATTACACAGAAAAGTCAAACCCT 538
      QY 121 GAGTTTTCATGATACGGAATCCAGCCATTTCTATTCATATT 164
      Db 537 TGGTTTTAAATGAAATCCCTTCATCCACCTGTGTCAAAAT 494

RESULT 3
CA439981/c
LOCUS
DEFINITION
  UI-H-D10-auf-1-07-0-UI.s1 NCI CGAP_D10 Homo sapiens cDNA clone
  UI-H-D10-auf-1-07-0-UI 3', mRNA sequence.
ACCESSION
  CA439981
VERSION
  CA439981.1 GI:24804401
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 556)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs@mail.nih.gov
  Tissue Procurement: Dr. Jose Mercuende
  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu

```

The following repetitive elements were found in this cDNA
 sequence: 526-556...>at_richflow_complexity
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES

source

Location/Qualifiers

```
1..556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-D10-aux-1-07-0-UI"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP D10"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified Polylinker; Site 1: EcoR 1; Site 2: Not I; NCI CGAP D10 is a cDNA library containing the following tissue(s): A pool of lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)18 tail. The sequence tag for this library is ATACGCGTC.
TAG_LIB=UI-H-D10
TAG_TISSUE=lung with fibrosis
TAG_SEQ=ATACGCGTC"
130 a 106 c 117 g 203 t
```

BASE COUNT

ORIGIN

```
Query Match 17.9%; Score 38.8; DB 14; Length 556;
Best Local Similarity 52.5%; Pred. No. 18; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 AATCGGACAGTATGAATACAGCGGATTAATACAGGTAAGTCTATTACAAGGAAAAAC 60
Db 177 AACTGGGCGATTATACCTCTTTATTAATCAAGTTGAAATGCTACAAAGCATAAAC 118

QY 61 CTTTAAAGATATATGAAGTATATCTGTTTGCCACAGGGAATGGTAAATATGCC 120
Db 117 ACTTTGCAAAAATCCACAGGCGACAGTTGTTCATCAACAGAAAAAGTCAAAACCACT 58

QY 121 GAGTTTTCACATGAATAGCAATCCAGGCAATTCCTATTTCATA 162
Db 57 TGGTTTAAATGAAATCCITTCATCCACTGTTTCATA 16
```

RESULT 4

```
AZ525328
LOCUS 600 bp DNA linear GSS 07-MAY-2001
DEFINITION 242pbG08 pb MEN #21 Plasmodium berghei genomic 3', genomic survey sequence.
ACCESSION AZ525328
VERSION AZ525328.1 GI:13966072
KEYWORDS GSS.
SOURCE Plasmodium berghei
ORGANISM Plasmodium berghei
REFERENCE 1 (bases 1 to 600)
AUTHORS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
COMMENT Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
```

Seq primer: M13(-20) forward
 Class: shotgun.

FEATURES

source

Location/Qualifiers

```
1..600
/organism="Plasmodium berghei"
/mol_type="genomic DNA"
/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
/db_xref="taxon:5821"
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
/clone_lib="PB MEN #21"
/note="Vector: pBluescript SK(+), phagemid excised from lambda ZAP; Site 1: EcoRV; Site 2: EcoRV; Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from contaminating host DNA by Hoechst Dye 33258-CsCl ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."
252 a 51 c 90 g 205 t 2 others
```

BASE COUNT

ORIGIN

```
Query Match 17.8%; Score 38.6; DB 28; Length 600;
Best Local Similarity 49.0%; Pred. No. 21;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 12 TATGAATACAGCGATTATACAGGTAAGTCTATTACAAGGAAAAACCTTTAAAGAAAT 71
Db 366 TAAATAATGCCAAATAACATCAAGGGTTAGTAGACAATAATAATCTATATCTATA 425

QY 72 AATATGAAGGTATTACCTTGTTCACAGGGAATGGTAAATATGCCGAGTTTTCAC 131
Db 426 AAGATAAGTGGATGAAATTTAATCAAAATGTTTTCAGTATATGCACATAATTTCGAA 485

QY 132 TGAATAGCGAATCCAGCCATTTCTTATTCATATTTGACTGGATGGCTGAATGCGATTTA 191
Db 486 GGCATGCGAATCAAGGATGGGAAATATAAANGCACTTAATTTGTTGATGATCTAA 545

QY 192 TAGATAATGACGATGAGATTTAATT 217
Db 546 TTGATGATGATATACAAATTTTGAAT 571
```

RESULT 5

```
BH146874
LOCUS 956 bp DNA linear GSS 27-AUG-2001
DEFINITION ENTQ15TR Entamoeba histolytica sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION BH146874
VERSION BH146874.1 GI:15302923
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 956)
AUTHORS Eukaryota; Entamoebidae; Entamoeba.
TITLE Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
JOURNAL Determination of clone end sequences from Entamoeba histolytica
COMMENT HM1:IMSS sheared DNA library (2001)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
```

Email: bjoftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Reverse
Class: shotgun

High quality sequence start: 47

High quality sequence stop: 458

Location/Qualifiers

FEATURES

source

1. 956
/organism="Entamoeba histolytica"

/mol_type="Genomic DNA"

/strain="HM1:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pBOS1; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

311 a 158 c 226 g 261 t

BASE COUNT

ORIGIN

Query Match 17.8%; Score 38.6; DB 28; Length 956;

Best Local Similarity 56.8%; Pred. No. 22;

Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 25 GATTATACAGGTAAAGTTATTACACGGAAAAACCTTTAAAGATAATATGAAGTA 84

DB 575 GATTGAATAGTTAAAGAAACCTCAAGAACGCTACCTGACGGATAATATGGAA 634

QY 85 TTACCTGTTGCCAACGGGAATGTAATATGCCGAGTGTTCCTCACTGAATACGAATC 144

DB 635 TTATTGTTTGTGAACAAGGAATTCCTCGAATTTCCAAAGTCTTTTACTCCAATGGAATA 694

QY 145 CAGCC 149

DB 695 CGGCC 699

RESULT 6

AA294909/c

LOCUS

DEFINITION

AA294909

VERSION

AA294909.1 GI:1947264

EST.

KEYWORDS

SOURCE

ORGANISM

Hom sapiens (human)

Hom sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 380)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fullmer, R.A., Bult

, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,

Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald

, L.M., Fitzhugh, W.M., Fitchman, J.L., Geoghagen, N.S., Glodek, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,

Kelly, J.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Utterback, T.R., Weidman, J.F., Li, Y.,

Small, K.V., Spriggs, T.A., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

Bednarik, D.P., Cao, L., Cepeda, M.A., Fischer, C., Hastings, G.A., He, W.W.,

Dimke, D., Feng, D.-F., Ferrie, A., Fisher, C., Kim, A.K., Kozak, D.L.,

Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Olsen, H., Raymond, L.,

Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Pannon

, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and

Venter, J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THC181566

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi.html>)

Seq primer: M13 Reverse

Location/Qualifiers

1. 380

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):190411"

/db_xref="taxon:9606"

/dev_stage="adult"

/clone_lib="Pancreas tumor I"

/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:

ECORI; Site_2: XhoI"

99 a 48 c 79 g 152 t 2 others

BASE COUNT

ORIGIN

Query Match 17.6%; Score 38.2; DB 9; Length 380;

Best Local Similarity 51.8%; Pred. No. 25;

Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGACAGTATGAATACAGCGATTATACAGAGTGAAGTTCAATCAACGGAAAAAC 60

DB 295 AACTGAGGATTATACCTCTTTATATAATCAAGTTGAAATGCTNACAAAGATAAAC 236

QY 61 CTTTAAAGATAATATGAAGGTATACCTTTGTCACACGGGAATGTTAAATATGCC 120

DB 235 ACTTCTGCAAAAATTCACAGGCACAGTTGTTCATTCACAGAAAAAGTCAAAACCACT 176

QY 121 GAGTTTTTCAGTATGAGGAATCCAGCCATTTCTTATTCATATT 164

DB 175 TGGTTTTTAATGAATATCTTCACATCCACCTGTGTTCAAAAT 132

RESULT 7

AA356543/c

LOCUS

DEFINITION

AA356543

VERSION

AA356543.1 GI:2008862

EST.

KEYWORDS

SOURCE

ORGANISM

Hom sapiens (human)

Hom sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 308)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fullmer, R.A., Bult

, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,

Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald

, L.M., Fitzhugh, W.M., Fitchman, J.L., Geoghagen, N.S., Glodek, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,

Kelly, J.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Utterback, T.R., Weidman, J.F., Li, Y.,

Small, K.V., Spriggs, T.A., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

Bednarik, D.P., Cao, L., Cepeda, M.A., Fischer, C., Hastings, G.A., He, W.W.,

Dimke, D., Feng, D.-F., Ferrie, A., Fisher, C., Kim, A.K., Kozak, D.L.,

Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Olsen, H., Raymond, L.,

Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hung, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 7566098
 Other ESTs: THC181566
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..308
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:158684"
 /db_xref="taxon:9606"
 /cell_type="T-lymphocyte"
 /clone_lib="Jurkat T-cells VI"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
 90 a 32 c 57 g 129 t

Query Match 17.3%; Score 37.6; DB 9; Length 308;
 Best Local Similarity 51.8%; Pred. No. 34;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 1 AATTCGGACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGAAAC 60
 Db 300 AACTGAGGCATTATACCTCTTTATATATGCAAGTTGAAATGCTAACAAAGCATTAAC 241
 QY 61 CTTTAAAGAAATATGAAGGTATTACCTTGTTCGCAAGGAAATGTAATATGCC 120
 Db 240 ACTTCGAAAATTCACAGGCAGAGTTGTTCAATCAACAGAAAAGTCAAAACCACT 181
 QY 121 GAGTTTTCACTGAATAGCAATCCAGCCATTCTTAATCATATT 164
 Db 180 TGGTTTTAAATGAAAATCCCTTCACATCCACTGTGTTCAAAAT 137

RESULT 8
 AA23582/c
 LOCUS EST511485 Human Gall bladder Homo sapiens cDNA 5' end similar to
 DEFINITION None. mRNA sequence.
 ACCESSION T32582
 VERSION T32582.1 GI:614680
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 312)
 REFERENCE Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M.-W., Clayton, R.A., Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D., FitzGerald, L.M., FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S.W., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S., Kelley, J.M., Klinek, K.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pellegrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.-J., Dimke, D., Feng, P., Ferrie, A., Fischer, C., Hastings, G.A., He, W.-W., Hu, J.-S., Greene, J.M., Gruber, J., Hudson, P., Kim, A., Kozak, D.L., Kunsch, C., Ji, H., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.-F., Wing, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
 Initial Assessment of Human Gene Diversity and Expression Patterns Based Upon 83 Million Basepairs of cDNA Sequence
 Nature 377, 3-174 (1995)
 96026280
 7566098
 Other ESTs: THC9475
 Contact: Venter, JC
 The Institute for Genomic Research
 932 Clopper Rd, Gaithersburg, MD 20878
 Tel: 3018699056
 Fax: 3018699423
 Email: tdbinfo@db.tigr.org
 For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database (tdbinfo@db.tigr.org)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..312
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):103165"
 /db_xref="taxon:9606"
 /clone_lib="Human Gall bladder"
 /note="Organ: gallbladder"
 80 a 36 c 64 g 132 t

Query Match 17.3%; Score 37.6; DB 14; Length 312;
 Best Local Similarity 51.8%; Pred. No. 34;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 1 AATTCGGACAGTATGAATACAGCGGATTAAATACAGGTAAGTTCAATACACGGAAAC 60
 Db 246 AACTGAGGCATTATACCTCTTTATATATGCAAGTTGAAATGCTAACAAAGCATTAAC 187
 QY 61 CTTTAAAGAAATATGAAGGTATTACCTTGTTCGCAAGGAAATGTAATATGCC 120
 Db 186 ACTTCGAAAATTCACAGGCAGAGTTGTTCAATCAACAGAAAAGTCAAAACCACT 127
 QY 121 GAGTTTTCACTGAATAGCAATCCAGCCATTCTTAATCATATT 164
 Db 126 TGGTTTTAAATGAAAATCCCTTCACATCCACTGTGTTCAAAAT 83

RESULT 9
 AA294908/c
 LOCUS EST100071 Pancreas tumor I Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION AA294908
 ACCESSION AA294908
 VERSION AA294908.1 GI:1947263
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 324)
 REFERENCE Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., FitzGerald, L.M., FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

RESULT 9
 AA294908/c
 LOCUS EST100071 Pancreas tumor I Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION AA294908
 ACCESSION AA294908
 VERSION AA294908.1 GI:1947263
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 324)
 REFERENCE Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., FitzGerald, L.M., FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Fertie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THC181566

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse

Location/Qualifiers

1. .324

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):190410"

/db_xref="taxon:9606"

/dev_stage="adult"

/clone_lib="Pancreas tumor 1"

/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 90 a 35 c 64 g 133 t 2 others

ORIGIN

Query Match 17.3%; Score 37.6; DB 9; Length 324;

Best Local Similarity 51.8%; Pred. No. 34;

Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATGAATACACCGGATTATATACAGGTAAAGTTCAATTCATTAACCGGAAAAAC 60

DB 295 AACTGAGCAATTAATACCTTTTATATATGCAAGTTGAATCTTCAACAAAGCATTAAC 236

QY 61 CTTTAAAGAAATATGAAGTATTACCTTTGTTGCCAACGGGAATGGTAATATGCC 120

DB 235 ACTTCTGCAAAAATTCACAGGACAGTGTTCATTCAACAGAAAAGTCAAAACCACT 176

QY 121 GAGTTTTCACATGAATACGGAATCCAGCAATTCATTATCATTT 164

DB 175 TCGTTTAAATGAATCCCTTCATCCACCTGTGTCATAAT 132

RESULT 10

R69154/c

LOCUS

DEFINITION

IMAGE:142136 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Werck EST Project

Unpublished

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

Insert Size: 966

High quality sequence stops: 245

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 966 Std Error: 0.00

Seq primer: Promega -21ml3

High quality sequence stop: 245.

Location/Qualifiers

1. .340

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:550897"

/db_xref="taxon:9606"

/clone="IMAGE:142136"

/sex="Female"

/dev_stage="placenta obtained at birth (full term)"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares placenta Nb2Hp"

/note="Organ: placenta; Vector: p773D (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5',

AACTCGAAGAAATCGCGCGAGAAATTTTTTTTTTTTTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified p773 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 85 a 47 c 74 g 132 t 2 others

ORIGIN

Query Match 17.3%; Score 37.6; DB 14; Length 340;

Best Local Similarity 51.8%; Pred. No. 35;

Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGACAGTATGAATACACCGGATTATATACAGGTAAAGTTCAATTCATTAACCGGAAAAAC 60

DB 215 AACTGAGCAATTAATACCTTTTATATATGCAAGTTGAATCTTCAACAAAGCATTAAC 156

QY 61 CTTTAAAGAAATATGAAGTATTACCTTTGTTGCCAACGGGAATGGTAATATGCC 120

DB 155 ACTTCTGCAAAAATTCACAGGACAGTGTTCATTCAACAGAAAAGTCAAAACCACT 96

QY 121 GAGTTTTCACATGAATACGGAATCCAGCAATTCATTATCATTT 164

DB 95 TCGTTTAAATGAATCCCTTCATCCACCTGTGTCATAAT 52

RESULT 11

BI017065

LOCUS

DEFINITION

PM3-ET0272-300301-002-b06 ET0272 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BI017065

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 200202663
 MEDLINE
 10737800
 PUBMED
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PM3&t2=PM3-ET0272-300301-002-b06&t3=2001-03-30&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 5
 High quality sequence stop: 402.

FEATURES

source
 1..402
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="ET0272"

/note="Organ: lung_tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 163 a 79 c 58 g 101 t 1 others

Query Match 17.3%; Score 37.6; DB 12; Length 402;
 Best Local Similarity 51.8%; Pred. No. 35;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 1 AATTGGCAGTATGAATACAGCGGATTATACAGGTAAAGTTCAATTCATTACACGGAAAC 60
 DB 120 AACTCGCAGTATGAATACCTCTTTATATATGCAAGTTGAAATGCTAACAAAGCATAAAC 179
 QY 61 CTTTAAAGATAATATGAAGGTATACCTTGTTCGCAAGGAAATGTTAAATATGCC 120
 DB 180 ACTTCGCAAAATTCACAGGCACAGTTGTTCAATCAACAGAAAGTCAAAACCACT 239
 QY 121 GAGTTTTTCACTAGTAGCAATCCAGCCATTTCTATTATATTT 164
 DB 240 TGGTTTTTAATGAAATCCCTTCACATCCACCTGTGTTCAAAAT 283

RESULT 12
 BF446491/c
 LOCUS
 DEFINITION
 mRNA sequence.
 ACCESSION
 BF446491
 VERSION
 BF446491.1 GI:11511629
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 407)
 REFERENCE
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 405.

FEATURES

source
 1..407
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:3705125"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Lu24"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 120 a 46 c 77 g 164 t

Query Match 17.3%; Score 37.6; DB 10; Length 407;
 Best Local Similarity 51.8%; Pred. No. 35;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 1 AATTGGCAGTATGAATACAGCGGATTATACAGGTAAAGTTCAATTCATTACACGGAAAC 60
 DB 392 AACTGAGGCATTATACCTCTTTATATATGCAAGTTGAAATGCTAACAAAGCATAAAC 333
 QY 61 CTTTAAAGATAATATGAAGGTATACCTTGTTCGCAAGGAAATGTTAAATATGCC 120
 DB 332 ACTTCGCAAAATTCACAGGCACAGTTGTTCAATCAACAGAAAGTCAAAACCACT 273
 QY 121 GAGTTTTTCACTAGTAGCAATCCAGCCATTTCTATTATATTT 164
 DB 272 TGGTTTTTAATGAAATCCCTTCACATCCACCTGTGTTCAAAAT 229

RESULT 13
 BM755155/c
 LOCUS
 DEFINITION
 K-EST0032842 S1LSNU1 Homo sapiens cDNA clone S1LSNU1-7-C07 5', mRNA sequence.
 ACCESSION
 BM755155
 VERSION
 BM755155.1 GI:19084773
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 435)
 REFERENCE
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished
 Contact: Kim YS
 Genom Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470

Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 7 row: C column: 07
High quality sequence stop: 435.
Location/Qualifiers

FEATURES

source

1. 435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-7-C07"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"
/lab_host="Top10P"
/clone_lib="S11SNUI"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

BASE COUNT 105 a 81 c 91 g 158 t

ORIGIN

Query Match 17.3%; Score 37.6; DB 12; Length 435;
Best Local Similarity 51.8%; Pred. No. 36;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 1 AATTCGGACATGATACACGGGATTAATACAGGTAGTTCATTACACGGAAAAAC 60
DB 195 AACTGCGAAATATCCACAGGCACAGTGTTCATTACACAGAAAAAGTCATAAAC 136
QY 61 CTTTAAAGATAATATGAAGGTATTACCTTTTGGCAACGGGAATGGTAAATATCCCC 120
DB 135 ACTTCTGCAAAAATCCACAGGCACAGTGTTCATTACACAGAAAAAGTCATAAAC 76
QY 121 GAGTTTTCACATGATACGGAATCCAGGCATTTCTATTCATATT 164
DB 75 TGGTTTTTAAATGAAATCCTTCACATCCACCTGTGTCAAAAT 32

RESULT 14

N57411/c

LOCUS N57411 437 bp mRNA linear EST 22-FEB-1996
DEFINITION yw82c07.r1 Soares_placenta.8to9weeks.2NDHP8to9w Homo sapiens cDNA
clone IMAGE:258732 5', mRNA sequence.

ACCESSION

N57411

VERSION N57411.1 GI:1201301

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 437)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Merck EST Project

UNPUBLISHED

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through INL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: F7

High quality sequence stop: 344.

FEATURES

source

1. 437
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3888438"
/db_xref="taxon:9606"
/clone="IMAGE:258732"
/dev_stages="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_placenta.8to9weeks.2NDHP8to9w"
/note="Organ: placenta; Vector: p77T1D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCGATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77T3 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 117 a 60 c 87 g 173 t

ORIGIN

Query Match 17.3%; Score 37.6; DB 14; Length 437;
Best Local Similarity 51.8%; Pred. No. 36;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 1 AATTCGGACATGATACACGGGATTAATACAGGTAGTTCATTACACGGAAAAAC 60
DB 310 AACTGCGAAATATCCACAGGCACAGTGTTCATTACACAGAAAAAGTCATAAAC 251
QY 61 CTTTAAAGATAATATGAAGGTATTACCTTTTGGCAACGGGAATGGTAAATATGCC 120
DB 250 ACTTCTGCAAAAATCCACAGGCACAGTGTTCATTACACAGAAAAAGTCATAAAC 191
QY 121 GAGTTTTCACATGATAGGAATCCAGGCATTTCTATTCATATT 164
DB 190 TGGTTTTTAAATGAAATCCTTCACATCCACCTGTGTCAAAAT 147

RESULT 15

H48409/c

LOCUS

H48409

DEFINITION

yr32f06.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone

IMAGE:207011 5', mRNA sequence.

ACCESSION

H48409

VERSION

H48409.1 GI:986796

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 441)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Merck EST Project

UNPUBLISHED

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 895

High quality sequence stops: 308
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 895 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 308.

FEATURES
 Location/Qualifiers
 1..441
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3776142"
 /db_xref="taxon:9606"
 /clone="IMAGE:207011"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INFLS"
 /note="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5', AACTGGAAGATTAATAAGACCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTV73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."
 BASE COUNT 103 a 80 c 96 g 160 t 2 others
 ORIGIN

Query Match 17.3%; Score 37.6; DB 14; Length 441;
 Best Local Similarity 51.8%; Pred. No. 36;
 Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 AATTCGGACAGTATCAATACAGCGGATTAAATACAGGTAAGTTCATTACACGGGAAAAAC 60
 |||
 Db 186 AACTGAGGCATTATACCTCTTTATATATGCAAGTTGAAATGCTACAAAGCATTAAC 127
 |||
 QY 61 CTTTAAAGATAATATGAAGGTATTACCTTGTTCACACGGGAATGGTAAATATGCC 120
 |||
 Db 126 ACTTCTGCAAAAATTCACAGGCACAGTTGTTCAATTCACAGAAAAGTCAAAACCACT 67
 |||
 QY 121 GAGTTTTTCACTGAATACGGAATCCAGGCATTCTATTATATT 164
 |||
 Db 66 TGGTTTTTAAATGAAAAATCCTTTCATCCACCTGCTGTTCAAAAT 23
 |||

Search completed: November 15, 2003, 08:02:54
 Job time : 1451.63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 15, 2003, 00:06:39 ; Search time 33.7995 Seconds
(without alignments)
2833.774 Million cell updates/sec

Title: US-09-928-457-99
Perfect score: 217
Sequence: 1 AATTCGGACAGTATGATAC.....ATGACGATGAGATTAAATT 217

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.6	17.3	3826	4	US-09-220-132-51
C 2	34.2	15.8	2047	3	US-08-836-261A-1
C 3	32.4	14.9	1830121	4	US-09-557-884-1
C 4	32.4	14.9	1830121	4	US-09-643-990A-1
C 5	32	14.7	1684976	4	US-08-916-421B-1
C 6	31.6	14.6	1478	1	US-08-090-523-27
C 7	31.6	14.6	1478	1	US-08-398-627-27
C 8	31.6	14.6	1478	1	US-08-406-857-1
C 9	31.6	14.6	1478	1	US-08-596-024-3
C 10	31.6	14.6	1478	3	US-09-020-818-3
C 11	31.6	14.6	1478	3	US-08-907-740-3
C 12	31.6	14.6	1478	4	US-09-797-467-3
C 13	31.6	14.6	1478	4	US-08-399-023-27
C 14	31.6	14.6	1478	5	PCT-US94-07072-1
C 15	31.2	14.4	3030	4	US-09-328-352-4115
C 16	30.6	14.1	1184	1	US-08-362-511A-1
C 17	30.6	14.1	1239	1	US-08-362-511A-1
C 18	30.6	14.1	1684976	4	US-08-916-421B-1
C 19	30.2	13.9	7418	4	US-09-620-312D-365
C 20	30	13.8	2142	4	US-09-107-532A-905
C 21	30	13.8	148567	4	US-09-801-876B-3
C 22	29.6	13.6	1172	4	US-08-969-046-5
C 23	29.6	13.6	2255	3	US-08-714-918-105
C 24	29.6	13.6	2255	3	US-09-265-315-105
C 25	28.6	13.6	2255	3	US-09-265-315-105
C 26	29.6	13.6	2255	3	US-09-266-417-105
C 27	29.4	13.5	3600	3	US-08-894-731-1

C 28	29.2	13.5	419	3	US-09-030-607-201	Sequence 201, App
C 29	29.2	13.5	419	4	US-09-439-313-201	Sequence 201, App
C 30	29.2	13.5	419	4	US-09-352-616A-201	Sequence 201, App
C 31	29.2	13.5	419	4	US-09-232-149A-201	Sequence 201, App
C 32	29.2	13.5	1230025	4	US-09-198-452A-1	Sequence 73, Appli
C 33	28.8	13.3	597	4	US-09-712-016-73	Sequence 73, Appli
C 34	28.6	13.2	1977	2	US-08-825-558-3	Sequence 3, Appli
C 35	28.6	13.2	1977	4	US-09-312-611-3	Sequence 3, Appli
C 36	28.6	13.2	2369	1	US-07-797-556-1	Sequence 1, Appli
C 37	28.6	13.2	2369	1	US-08-308-881-1	Sequence 1, Appli
C 38	28.6	13.2	2369	2	US-09-058-263-1	Sequence 1, Appli
C 39	28.6	13.2	2369	2	US-09-059-099-1	Sequence 1, Appli
C 40	28.6	13.2	2369	3	US-09-058-262-1	Sequence 1, Appli
C 41	28.6	13.2	2369	4	US-09-455-962-1	Sequence 1, Appli
C 42	28.6	13.2	2369	5	PCT-US95-06530-1	Sequence 1, Appli
C 43	28.6	13.2	2754	2	US-08-825-558-5	Sequence 5, Appli
C 44	28.6	13.2	2754	4	US-09-312-611-5	Sequence 5, Appli
C 45	28.6	13.2	3085	3	US-08-795-473B-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-220-132-51/c
; Sequence 51, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 3826
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-220-132-51

Query Match	17.3%;	Score 37.6;	DB 4;	Length 3826;
Best Local Similarity	51.8%;	Pred. No. 0.041;		
Matches	85;	Conservative	0;	Mismatches 79; Indels 0; Gaps 0;
Qy	1	AATTCGGACAGTATGATACAGCGGATTATTAACAAGTAAAGTTTCATTACAACGGAAAAAC	60	
Db	3162	AACGTGAGGCAATTAATACCTCTTTATTAATAATGCAAGTTGAAATGCTTACAAAGCATAAAC	3103	
Qy	61	CTTTAAAGATATATGAAGGTATTACCTGTTGCCAACGGATGTAATATGCCC	120	
Db	3102	ACTTCGCAAAATTCACCAAGCAGTTGTTTCATTCAACAGAAAAGTCAAAACCACT	3043	
Qy	121	GAGTTTTTCAGTGAATAGGGAATCCAGCCATTTCATTTCATATT	164	
Db	3042	TGTTTTTAAATGAAATCTTCACATCCACCTGTGTTCAAAAT	2999	

RESULT 2

US-08-836-261A-1
; Sequence 1, Application US/08836261A
; Patent No. 6221582
; GENERAL INFORMATION:
; APPLICANT: GIESENDORF, BELINDA
; APPLICANT: CUINT, WILHELMUS
; APPLICANT: VAN DOORN, LEENDERT-JAN
; TITLE OF INVENTION: NEW POLYNUCLEIC ACID SEQUENCES FOR USE IN THE
; DETECTION AND DIFFERENTIATION OF PROKARYOTIC ORGANISMS

NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836.261A
FILING DATE: 25 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04264
FILING DATE: 30 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870171.9
FILING DATE: 28 Oct 1994
ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:005
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2047 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 864..2016
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 864..2013
US-08-836-261A-1

Query Match 15.8%; Score 34.2; DB 3; Length 2047;
Best Local Similarity 49.2%; Pred. No. 0.35;
Matches 90; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 34 AAGSTAAGTTTCATTACACAGGAAACCTTTTAAAGAATAATATGAAAGGTATTACCTTGT 93
Db 824 AATTTTATATAATTTTAAATAATAATTTCTAGGAAAAAATGCAAAGCATCATCTTA 883
QY 94 TTGCCAAGGGAATGTTAAATATGCCAGTTTTCCTGTAATAGCGAATCCAGCCATTT 153
Db 884 TAGGCAAGCCAAATGTTGGAAATCAAGTCCTTTTATAGAAATGCGAAGCAAGAAATAG 943
QY 154 CTATTTCATATTGACTGGATGCTCAATGTGGACTTTTATAGATAATGACGATGAAGATTT 213
Db 944 CTATTACAGTCATATTTCAGGTACACTAGATACAAATAAAGCAAAATTCATATTC 1003
QY 214 AAT 216
Db 1004 ATT 1006

RESULT 3
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-3-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 14.9%; Score 32.4; DB 4; Length 1830121;
Best Local Similarity 48.9%; Pred. No. 8.2;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 2 ATTCGACAGTATGAATACAGCGGATTAATACAGGTAACTTATACACGGAACACC 61
Db 761222 AATTCGTCAAGATGTATTAAACGATTTTATACCGTCATTTTTTAATGCTGTTATTGC 761281
QY 62 TTTTAAAGATAATATGAAAGGTATTACCTTGTTCGCAACGGAATGGTAAATATGCCCG 121
Db 761282 AGAAAAATTAATATGAGTCGTCGCAACTTTTCTATTGAACTTTTGAAGAGTAA 761341
QY 122 AGTTTTCCTACTGAATAGGAATCCAGCCATTTCTATTCTATTTGACTGGATGGCTGA 179
Db 761342 AGACTAGTTTTTCACTGCAACTTTCGAGGTTTATCCAGAAGTTAAATTACAAGGTTTA 761399

RESULT 4
US-09-643-990A-1
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186PIC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 14.9%; Score 32.4; DB 4; Length 1830121;
Best Local Similarity 48.9%; Pred. No. 8.2;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 2 ATTCCGACAGATGATACACGGGATTATACAGGTAAGTTTCATTACACGGAAACCC 61
DB 761222 AATTCGTCAGATGATTAACGATTATTACCAGTCATTTTAAATGCTGTTATGC 761281

QY 62 TTTAAGAATAATATGAAGGTATPACCTGTTTCCCAACGGGAATGTAATATGCCCG 121
DB 761282 AGAAAAATTAATATTCAGGTCGTCACATTTTGCTATTGAACATTTTGAGAGGTAA 761341

QY 122 AGTTTTCACATGATGATCAACGATTCATTCATTTCTATTTGATGCTGCTGA 179
DB 761342 AGACTTAGTTTCTAGTCACTTTGAGGTTTATCCAGAGTTAAATACAAAGTTTA 761399

RESULT 5
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916.421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature

```

; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1
Query Match 14.7%; Score 32; DB 4; Length 1664976;
Best Local Similarity 65.3%; Pred. No. 11;
Matches 47; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 24 GGATTATACAGGTAGTTCATTACACGGAAACCTTTAAAGCATATATGAAAGCT 83
Db 953339 GGATTACATTAGGTGATTCAATTAAGAAATTAAGAGTAAATATGGAAGGA 953280
QY 84 ATTACCTTGTTT 95
Db 953279 GAAAGCTTATTT 953268

RESULT 6
US-08-090-523-27/c
; Sequence 27, Application US/08090523
; Patent No. 5498830
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Enhanced Starch Biosynthesis
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Co. B4.F
; STREET: 700 Chesterfield Parkway No. 5498830th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/090,523
; FILING DATE: 19930712
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/709663
; FILING DATE: 07-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539763
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10559)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-7286
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-090-523-27
Query Match 14.6%; Score 31.6; DB 1; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 2 ATTCGACAGTATGATACGCGGATTATACAGGTAGTTCATTACACGGAAACCT 61
```

Db 615 AATCTTACATTAATAAATAACATATAAAATATCCTTTATTTTGTAAATTTTGAATATGT 556
QY 62 TTTAAAGAAATATGAAAGGTATTACCTTGTGTCACGGAATGGTAAATATGCCCG 121
Db 555 TATGTAAGTTATATTAAATTTTATCGATTTTAAAAAGTGAATGAAATTTTAGTG 496
QY 122 AGTTTTCCTGTAAGCAATCCAGCCATTTCTATTCTATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

RESULT 7
US-08-398-627-27/c
; Sequence 27, Application US/08398627
; Patent No. 5608149
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Enhanced Starch Biosynthesis
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Co. B4f
; STREET: 700 Chesterfield Parkway No. 5608149th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,627
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/090,523
; FILING DATE: 12-JUL-1993
; APPLICATION NUMBER: US 07/709663
; FILING DATE: 07-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539763
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10559)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-7286
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-398-627-27

Query Match 14.6%; Score 31.6; DB 1; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGGACAGTATGAATACAGCGGATTAATACAGGTAAGTTCAATACACGGAATAACC 61
Db 615 AATCTTACATTAATAAATAACATATAAAATATCCTTTATTTTGTAAATTTTGAATATGT 556
QY 62 TTTAAAGAAATATGAAAGGTATTACCTTGTGTCACGGAATGGTAAATATGCCCG 121
Db 555 TATGTAAGTTATATTAAATTTTATCGATTTTAAAAAGTGAATGAAATTTTAGTG 496

QY 122 AGTTTTCCTGTAAGCAATCCAGCCATTTCTATTCTATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

RESULT 8
US-08-406-857-1/c
; Sequence 1, Application US/08406857
; Patent No. 5608150
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; TITLE OF INVENTION: Fruit Specific Promoters
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, B4f
; STREET: 700 Chesterfield Parkway No. 5608150th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,857
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07072
; FILING DATE: 27-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/090,523
; FILING DATE: 12-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10655)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-7286
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-406-857-1

Query Match 14.6%; Score 31.6; DB 1; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGGACAGTATGAATACAGCGGATTAATACAGGTAAGTTCAATACACGGAATAACC 61
Db 615 AATCTTACATTAATAAATAACATATAAAATATCCTTTATTTTGTAAATTTTGAATATGT 556
QY 62 TTTAAAGAAATATGAAAGGTATTACCTTGTGTCACGGAATGGTAAATATGCCCG 121
Db 555 TATGTAAGTTATATTAAATTTTATCGATTTTAAAAAGTGAATGAAATTTTAGTG 496
QY 122 AGTTTTCCTGTAAGCAATCCAGCCATTTCTATTCTATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

RESULT 9
US-08-596-024-3/c
; Sequence 3, Application US/08596024
; Patent No. 5716837
; GENERAL INFORMATION:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/596,024
FILING DATE: 06-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Patterson, Melinda L.
REGISTRATION NUMBER: 33,062
REFERENCE/DOCKET NUMBER: M08171/PAT/HUN
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1478 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-907-740-3

Query Match 14.6%; Score 31.6; DB 3; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGGACAGTATGAATACAGCGGATTATACAGGTAAGTTCATTACACGGAAAAACC 61
Db 615 AATCTTACATTAATAAATAAACATATAAAAAATATCCTTTATTTTGTAAATTTGATATGT 556
QY 62 TTAAAGAAATATATGAAGGTATTACCTTGTTCGCCAAGGGAATGTAATAATATGCCCCG 121
Db 555 TAATGAAGTTATATTAATAATTTATCGATTTTAAAGTAAGTAATGAATTTAGT 496
QY 122 AGTTTCTACTGAATAGCAATCCAGCCATTTCTATTCATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

RESULT 12
US-09-797-467-3/c
Sequence 3, Application US/09797467
Patent No. 6476295
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: DeWeerd, Jan W.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Welton, Marcia L.
TITLE OF INVENTION: Expression of Sucrose Phosphorylase in Plants
FILE REFERENCE: 11899.0004.CNUS02 M08T:004--2
CURRENT APPLICATION NUMBER: US/09/797,467
CURRENT FILING DATE: 2001-03-01
PRIOR FILING DATE: US 09/020,818
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: US 08/596,024
PRIOR FILING DATE: 1997-02-06
PRIOR APPLICATION NUMBER: US 08/386,860
PRIOR FILING DATE: 1995-02-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.0
SEQ ID NO 3
LENGTH: 1478
TYPE: DNA
ORGANISM: Genomic
US-09-797-467-3

Query Match 14.6%; Score 31.6; DB 4; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGGACAGTATGAATACAGCGGATTATACAGGTAAGTTCATTACACGGAAAAACC 61
Db 615 AATCTTACATTAATAAATAAACATATAAAAAATATCCTTTATTTTGTAAATTTGATATGT 556
QY 62 TTAAAGAAATATGAAGGTATTACCTTGTTCGCCAAGGGAATGTAATAATATGCCCCG 121

Db 555 TATGTAAGTTATATTAATAATTTATCGATTTTAAAAAGTGAATTTAGT 496
QY 122 AGTTTCTACTGAATAGCAATCCAGCCATTTCTATTCATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

RESULT 13
US-08-399-023-27/c
Sequence 27, Application US/08399023
Patent No. 6538179
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Stark, David M.
TITLE OF INVENTION: Enhanced Starch Biosynthesis
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Co. BB4F
STREET: 700 Chesterfield Parkway No. 6538179th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,023
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/090,523
FILING DATE:
APPLICATION NUMBER: US 07/709663
FILING DATE: 07-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539763
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10559)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-7286
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1478 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-399-023-27

Query Match 14.6%; Score 31.6; DB 4; Length 1478;
Best Local Similarity 49.4%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGGACAGTATGAATACAGCGGATTATACAGGTAAGTTCATTACACGGAAAAACC 61
Db 615 AATCTTACATTAATAAATAAACATATAAAAAATATCCTTTATTTTGTAAATTTGATATGT 556
QY 62 TTAAAGAAATATGAAGGTATTACCTTGTTCGCCAAGGGAATGTAATAATATGCCCCG 121
Db 555 TAATGAAGTTATATTAATAATTTATCGATTTTAAAGTAAGTAATGAATTTAGT 496
QY 122 AGTTTCTACTGAATAGCAATCCAGCCATTTCTATTCATATTTGA 167
Db 495 AGAGAAACAATGAACGAAGACGCAACTTGATATATATACCTGA 450

QY 72 AATATGAAAGGTATTACCTTGTTT 95
Db 2385 TTTATATATAATCTTACATTTT 2362

Search completed: November 15, 2003, 08:09:51
Job time : 42.7995 secs

RESULT 14
PCT-US94-07072-1/c
; Sequence 1, Application PC/TUS9407072
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Fruit Specific Promoters
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07072
; FILING DATE: 27-JUN-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/090523
; FILING DATE: 12-JUL-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US94-07072-1

Query Match 14.6%; Score 31.6; DB 5; Length 1478;
Best Local Similarity 49.4%; Pred. NO. 1.9;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2 ATTCGACAGTATGATACAGCGGATTATACAGGTAAGTTCATTACACGGAAACC 61
Db 615 AATCTTACATTAATAAATAACATATAAAATATCCTTTATTTTGTGAATGT 556

QY 62 TTAAAGAAATATATGAAGTATTACCTTGTTCACCGGAATGGTAAATATGCCCG 121
Db 555 TATGTAAGTATATATAAAATTTATCGATTTTATAAAGTGAATTGAAATTTAGTG 496

QY 122 AGTTTTCATCAGTAGGATCCAGCCATTCCTTCATCATATTGA 167
Db 495 AGAGAAACAATGACGACGAAGACGACCACTTGATATATATACCTGA 450

RESULT 15
US-09-328-352-4115/c
; Sequence 4115, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4115
; LENGTH: 3030
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4115

Query Match 14.4%; Score 31.2; DB 4; Length 3030;
Best Local Similarity 60.7%; Pred. NO. 3.1;
Matches 51; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 12 TATGATACAGCGGATTATACAGGTAAGTTCATTACACGGAAACCCTTTAAGAAT 71
Db 2445 TATGAATCATCTATGTATATAAAATTAATTTTATAGACTTACAACTCTTTAAACT 2386

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	217	100.0	217	10	US-09-928-457-99	Sequence 99, Appl
C 2	37.6	17.3	507	11	US-09-918-995-22577	Sequence 22577, A
C 3	37.6	17.3	602	10	US-09-919-580-730	Sequence 730, App
C 4	37.6	17.3	2988	9	US-09-925-322-595	Sequence 295, App
C 5	37.6	17.3	3920	14	US-10-084-817-133	Sequence 133, App
C 6	37.6	17.3	5144	11	US-09-919-039-70	Sequence 70, Appl
7	36	16.6	2000	10	US-09-938-842A-3712	Sequence 3712, Ap
8	36	16.6	2003	10	US-09-987-576-306	Sequence 306, App
9	36	16.6	2004	10	US-09-987-576-253	Sequence 253, App
10	36	16.6	640681	10	US-09-790-588-1	Sequence 1, Appli
C 11	34.8	15.0	17421	12	US-10-240-453-56	Sequence 56, Appl
C 12	34.8	15.0	17421	14	US-10-239-776-54	Sequence 54, Appl
C 13	33	15.2	334	10	US-09-928-457-19	Sequence 19, Appl
C 14	33	15.2	495	11	US-09-918-995-22573	Sequence 22573, A
C 15	32.8	15.1	18434	12	US-10-311-455-1979	Sequence 1979, Ap
16	32.6	15.0	15881	12	US-10-311-455-234	Sequence 234, App

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-730

Query Match      17.3%; Score 37.6; DB 10; Length 602;
Best Local Similarity 51.8%; Pred. No. 1.8;
Matches      85; Conservative      0; Mismatches      79; Indels      0; Gaps      0;

QY      1 AATTCGGACAGTATGAATACACGCGGATTAATACAGGTAAGTTCATTACACGGAATAAC 60
Db      1 AATTCGGACAGTATGAATACACGCGGATTAATACAGGTAAGTTCATTACACGGAATAAC 60
QY      417 AACTGAGGCATTATACCTCTTTATATAATGCAAGTTGAAATGCTAAACAAGCATAAAC 358
Db      417 AACTGAGGCATTATACCTCTTTATATAATGCAAGTTGAAATGCTAAACAAGCATAAAC 358
QY      61 CTTTAAAGAAATAATGAAAGGTATTACCTTGTTGCCAAGCGGAATGTTAAATATGCC 120
Db      61 CTTTAAAGAAATAATGAAAGGTATTACCTTGTTGCCAAGCGGAATGTTAAATATGCC 120
QY      357 ACTTCTGCAGAAATTCACAGGCGACAGTTGTTCAATCAACAGAAAAAGTCAAAACCACT 298
Db      357 ACTTCTGCAGAAATTCACAGGCGACAGTTGTTCAATCAACAGAAAAAGTCAAAACCACT 298
QY      121 GAGTTTTCCTGAATAGCGAATCCAGGCATTTCTATTTCATTT 164
Db      121 GAGTTTTCCTGAATAGCGAATCCAGGCATTTCTATTTCATTT 164
QY      297 TGGTTTTTAAATGAAATCCITTCACATCCACCTGTGTCAAAT 254
Db      297 TGGTTTTTAAATGAAATCCITTCACATCCACCTGTGTCAAAT 254

RESULT 4
US-09-925-302-295/c
; Sequence 295, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 295
; LENGTH: 2998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (195)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2967)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2971)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2977)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2981)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-295

Query Match      17.3%; Score 37.6; DB 9; Length 2998;
Best Local Similarity 51.8%; Pred. No. 3.3;
Matches      85; Conservative      0; Mismatches      79; Indels      0; Gaps      0;

QY      1 AATTCGGACAGTATGAATACACGCGGATTAATACAGGTAAGTTCATTACACGGAATAAC 60
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-22577/c
; Sequence 22577, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22577
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(507)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22577

Query Match      17.3%; Score 37.6; DB 11; Length 507;
Best Local Similarity 51.8%; Pred. No. 1.7;
Matches      85; Conservative      0; Mismatches      79; Indels      0; Gaps      0;

QY      1 AATTCGGACAGTATGAATACACGCGGATTAATACAGGTAAGTTCATTACACGGAATAAC 60
Db      1 AATTCGGACAGTATGAATACACGCGGATTAATACAGGTAAGTTCATTACACGGAATAAC 60
QY      497 AACTGAGGCATTATACCTCTTTATATAATGCAAGTTGAAATGCTAAACAAGCATAAAC 438
Db      497 AACTGAGGCATTATACCTCTTTATATAATGCAAGTTGAAATGCTAAACAAGCATAAAC 438
QY      61 CTTTAAAGAAATAATGAAAGGTATTACCTTGTTGCCAAGCGGAATGTTAAATATGCC 120
Db      61 CTTTAAAGAAATAATGAAAGGTATTACCTTGTTGCCAAGCGGAATGTTAAATATGCC 120
QY      437 GCTTCTGCAAAAATTCACAGGCGACAGTTGTTCAATCAACAGAAAAAGTCAAAACCACT 378
Db      437 GCTTCTGCAAAAATTCACAGGCGACAGTTGTTCAATCAACAGAAAAAGTCAAAACCACT 378
QY      121 GAGTTTTCCTGAATAGCGAATCCAGGCATTTCTATTTCATTT 164
Db      121 GAGTTTTCCTGAATAGCGAATCCAGGCATTTCTATTTCATTT 164
QY      377 TGGTTTTTAAATGAAATCCITTCACATCCACCTGTGTCAAAT 334
Db      377 TGGTTTTTAAATGAAATCCITTCACATCCACCTGTGTCAAAT 334

RESULT 3
US-09-919-580-730/c
; Sequence 730, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 730
; LENGTH: 602
```

Db 2282 AACTGAGCAATTAATACCTCTTTTATAATGAAGTTGAAATGCTAAACAAGCATAAAC 2223
Qy 61 CTTTAAAGAATAATATGAAGGTATTACCTCTTTTCCAAACGGGAATGTTAAATATGCC 120
Db 2222 ACTTCTGCAAAATTCACCAAGGCACAGTTGTTCTATTCAACAGAAAAGTCAAAACCACT 2163
Qy 121 GAGTTTTCACCTGAATAGGAATCCAGGCATTTCTATTCAAT 164
Db 2162 TGGTTTTTAATGAAATCCITTCACATCCACCTGTGTTCAAAAT 2119

RESULT 5

US-10-084-817-133/c
; Sequence 133, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 133
; LENGTH: 3920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 407096.14
US-10-084-817-133

Query Match 17.3%; Score 37.6; DB 14; Length 3920;
Best Local Similarity 51.8%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 1 AATTCGACAGTATGAATACAGCGGATTAATACAGGTAAGTTCAATTAACAACGGAAAAAC 60
Db 3277 AACTGAGCAATTAATACCTCTTTTATAATGAAGTTGAAATGCTAAACAAGCATAAAC 3218
Qy 61 CTTTAAAGAATAATGAAGGTATTACCTCTTTTCCAAACGGGAATGTTAAATATGCC 120
Db 3217 ACTTCTGCAAAATTCACCAAGGCACAGTTGTTCTATTCAATTAACAAGAAAACCACT 3158
Qy 121 GAGTTTTCACCTGAATAGGAATCCAGGCATTTCTATTCAAT 164
Db 3157 TGGTTTTTAATGAAATCCITTCACATCCACCTGTGTTCAAAAT 3114

RESULT 6

US-09-919-039-70/c
; Sequence 70, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-03-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 70
; LENGTH: 5144
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 407096.2
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4537, 4540, 4590, 5133
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-039-70
Query Match 17.3%; Score 37.6; DB 11; Length 5144;
Best Local Similarity 51.8%; Pred. No. 4.1;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 1 AATTCGACAGTATGAATACAGCGGATTAATACAGGTAAGTTCAATTAACAACGGAAAAAC 60
Db 3277 AACTGAGCAATTAATACCTCTTTTATAATGAAGTTGAAATGCTAAACAAGCATAAAC 3218
Qy 61 CTTTAAAGAATAATGAAGGTATTACCTCTTTTCCAAACGGGAATGTTAAATATGCC 120
Db 3217 ACTTCTGCAAAATTCACCAAGGCACAGTTGTTCTATTCAATTAACAAGAAAACCACT 3158
Qy 121 GAGTTTTCACCTGAATAGGAATCCAGGCATTTCTATTCAAT 164
Db 3157 TGGTTTTTAATGAAATCCITTCACATCCACCTGTGTTCAAAAT 3114

RESULT 7

US-09-938-842A-3712
; Sequence 3712, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: S01P1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/254,647
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3712
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3712

Query Match 16.6%; Score 36; DB 10; Length 2000;
Best Local Similarity 51.9%; Pred. No. 7.8;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 29 AATCAAGGTAAGTTCAATTAACAACGGAAAAACCTTTAAAGATAATATGAAGGTATTAC 88
Db 13 AAACATCGAACGAAATTAATACGTAAAGCACCATAAAATGAGAAAATACACAACATAT 72
Qy 89 CTTGTTTGCAACGGGAATGTTAAATATGCCGAGTTTCTACTGAATAGCGAATCCAGC 148
Db 73 AGTGATTGACGATGCTTTTGACATATAAGTCTTACTCGTTAAGCAGATTATAATCTATT 132
Qy 149 CATTCTTATCATATTGACTGGATGGCTGAATGTC 184
Db 133 TTTTTCGATTGATCTTACGTTATGCAAAATATG 168

RESULT 8

US-09-887-576-306
; Sequence 306, Application US/09887576

```
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 306
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-306

Query Match      16.6%; Score 36; DB 10; Length 2003;
Best Local Similarity 51.9%; Pred. No. 7.8;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 29 AATCAAGTAAAGTTCATTACACGGAACCTTTAAAGAAATATATGAAAGGTATTAC 88
DB 13 AAACATCGACGAAATTAACGTAAAGCACCATAAAATGAGAAAATACACAACTAT 72
QY 89 CTTGTTGCCAACGGGAATGTAATATGCCCGAGTTTTTCACTGAATAGCAATCCAGC 148
DB 73 AGTGATTCACGATGCTTTTGACATATAAGTCTTACTCGTTAACAGATTATTAATCTATT 132
QY 149 CATTTCTATTATTTGACTGGATGCTGAATGTG 184
DB 133 TTTTTCGATTGTATCTTACGCTTATGCAAAATTATG 168

RESULT 9
US-09-887-576-253
; Sequence 253.. Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-253
```

```
Query Match      16.6%; Score 36; DB 10; Length 2004;
Best Local Similarity 51.9%; Pred. No. 7.8;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 29 AATCAAGTAAAGTTCATTACACGGAACCTTTAAAGAAATATATGAAAGGTATTAC 88
DB 13 AAACATCGACGAAATTAACGTAAAGCACCATAAAATGAGAAAATACACAACTAT 72
QY 89 CTTGTTGCCAACGGGAATGTAATATGCCCGAGTTTTTCACTGAATAGCAATCCAGC 148
DB 73 AGTGATTCACGATGCTTTTGACATATAAGTCTTACTCGTTAACAGATTATTAATCTATT 132
QY 149 CATTTCTATTATTTGACTGGATGCTGAATGTG 184
DB 133 TTTTTCGATTGTATCTTACGCTTATGCAAAATTATG 168

RESULT 10
US-09-790-988-1
; Sequence 1.. Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match      16.6%; Score 36; DB 10; Length 640681;
Best Local Similarity 55.6%; Pred. No. 67;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 27 TTAATACAAAGGTAAGTTCATTACACGGAACCTTTAAAGAAATATATGAAAGGTATT 86
DB 512832 TTAATTAAGCATCCAAAGAAATAGGAAATCTAAATGAATAGTATTATTATT 512891
QY 87 ACCTTGTTCGCAACGGGAATGTAATATGCCCGAGTTTTTCACTGAATAGCGAATCCA 146
DB 512892 AAACCTCTTCTTAAAGGACCTCTTAAACGACCAAAATTTGCTGGAATACCTAAACCA 512951
QY 147 GCCA 150
DB 512952 GTCA 512955

RESULT 11
US-10-240-453-56/c
; Sequence 56.. Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: By Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
```

```
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 56
; LENGTH: 17421
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (167, 1832..1833, 1836..1837, 1842, 1848, 1852, 1857,
; LOCATION: 1859..1860)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2953, 4024, 4630, 4733, 4767, 8161, 8189, 8200, 9142..9143)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9147, 9153, 9168, 9170, 9173, 10243, 10340, 10394, 11996)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (12002, 12005, 12012, 12102, 13715, 13717, 16077, 16097)
; US-10-240-453-56

Query Match      16.0%; Score 34.8; DB 12; Length 17421;
Best Local Similarity 65.4%; Pred. No. 37;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 AATACACGGGATTATACAGGTAGTTCATTACACCGGAAACCTTTAAAGAAATAATA 75
Db 6042 AATATAAAATATTAATACCATATATTTTATTTAACCATTAACCAATATAAAATAATA 5983

QY 76 TGAAGGTATTACCTTGT 93
Db 5982 TATAAATTATATCTTTT 5965

RESULT 12
US-10-239-676-54/c
; Sequence 54, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 54
; LENGTH: 17421
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (167, 1832..1833, 1836..1837, 1842, 1848, 1852, 1857, 1859..1860)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2953, 4024, 4630, 4733, 4767, 8161, 8189, 8200, 9142..9143)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9147, 9153, 9168, 9170, 9173, 10243, 10340, 10394, 11996)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (12002, 12005, 12012, 12102, 13715, 13717, 16077, 16097)
; US-10-239-676-54

Query Match      16.0%; Score 34.8; DB 14; Length 17421;
Best Local Similarity 65.4%; Pred. No. 37;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 AATACACGGGATTATACAGGTAGTTCATTACACCGGAAACCTTTAAAGAAATAATA 75
Db 6042 AATATAAAATATTAATACCATATATTTTATTTAACCATTAACCAATATAAAATAATA 5983

QY 76 TGAAGGTATTACCTTGT 93
Db 5982 TATAAATTATATCTTTT 5965

RESULT 13
US-09-928-457-19/c
; Sequence 19, Application US/09928457
; Patent No. US20020164603A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Neisseria meningitidis
; STRAIN: Z2491
; US-09-928-457-19

Query Match      15.2%; Score 33; DB 10; Length 334;
Best Local Similarity 51.0%; Pred. No. 26;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 35 AGTAAAGTTCATACACCGGAAACCTTTAAAGAAATAATAATGAAAGGTATTACCTTGT 94
Db 327 AGGTGCTATTATGTTTCGAATAATGCTATAACTATAATATAGAAAACAATATTACACTAGG 268

QY 95 TGCACACGGGAATGGTAAATATGCCGAGTTTTTCACTGAATAGCGAATCCAGCCATTTC 154
Db 267 TCGGTATAGGATGCTATTATTTCATATTTTCTGATGAAATGTGGTAGCGGATTGG 208
```

US-10-311-455-1979
Query Match 15.1%; Score 32.8; DB 12; Length 18434;
Best Local Similarity 50.6%; Pred. No. 1.3e+02;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 1 AATTCGGACAGTATGAATACAGCGGATTAAATACAGGTAAAGTTTCATTACACGGAAAAAC 60
DB 2860 AAATACGCCACTACACTCCAACTAAATACAAAAAAACCCCTATCTCAAAAAAATAA 2801
QY 61 CTTTAAAGATAAATATGAAGGTATTACCTTGTTCACACGGGAATGTTAAATATATGCC 120
DB 2800 AAAAAAATTAATAAATAATATATACCAATAAATAAATAAATAAATAAATAAATAA 2741
QY 121 GAGTTTTTCACTGAATAGCGGATCCAGCCATTCTTA 156
DB 2740 TAAATATTTTCAACACATTACAAATCAACCAATCTCAA 2705
Search completed: November 15, 2003, 08:32:49
Job time : 125.983 secs

QY 155 TATTCATATTTGACTGGATGCTGAATGTGGAC 187
DB 207 TACGATTAACTCAATGTACGAGTAATTTGAC 175
RESULT 14
US-09-918-995-22573/c
; Sequence 22573, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22573
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(495)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22573

Query Match 15.2%; Score 33; DB 11; Length 495;
Best Local Similarity 49.7%; Pred. No. 30;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 1 AATTCGGACAGTATGAATACAGCGGATTAAATACAGGTAAAGTTTCATTACACGGAAAAAC 60
DB 478 AACTGAGGCATTAAATACCTCTTTTATAATATGCAAGTTGAATGCTAACAAAGCATAAAC 419
QY 61 CTTTAAAGATAAATATGAAGGTATTACCTTGTTCACACGGGAATGTTAAATATATGCC 120
DB 418 ACTTCTGCAAAAATTCACAAAGGCACAGTTGTTTCATTCAACAGAAAAAGTCAAAACCACT 359
QY 121 GAGTTTTTCACTGAATAGCGGATCCAGCCATTTCATTATATTTGACT 169
DB 358 TGGTTTTTAAATGAAATCCTTCACCTTCACCTTGTTCAAAATATTAAT 310

QY 155 TATTCATATTTGACTGGATGCTGAATGTGGAC 187
DB 207 TACGATTAACTCAATGTACGAGTAATTTGAC 175
RESULT 15
US-10-311-455-1979/c
; Sequence 1979, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1979
; LENGTH: 18434
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)